



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 129485

TO: Mary Mosher
Location: REM/3D25/3C18
Art Unit : 1648
Monday, August 23, 2004

Case Serial Number: 09/851410

From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: (571) 272-2523
toby.port@uspto.gov

Search Notes

Blank sheet

STIC-Biotech/ChemLib

129485

From: Mosher, Mary
Sent: Tuesday, August 10, 2004 8:17 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search 09/851410

Please translate & search AA encoded by:

nt 5147-7129 of SEQ 6

nt 5117-7096 of SEQ 10

And search SEQ 8.

Please do regular search, oligo search for the seq 6 & 10 translation products, and interference search. Thanks.

Mary Mosher
AU 1648
571-272-0906
Office Rem 3D25
Mail Rem 3C18

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

Blank sheet

RESULT 2
AR002160
LOCUS AR002160 2049 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 3 from patent US 5741490.
ACCESSION AR002160
VERSION AR002160.1 GI:3963714
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
UNCLASSIFIED.
REFERENCE 1 (bases 1 to 2049)
AUTHORS Reyes, G.R., Bradley, D.W., Twu, J.-S., Purdy, M.A., Tam, A.W.,
Krawczynski, K.Z. and Yarbough, P.D.
TITLE Hepatitis B virus vaccine and method
JOURNAL Patent: US 5741490-A 3 21-APR-1998;
FEATURES
Location/Qualifiers
1..2049
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 1983; DB 6; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGCTCGGCTATTTTGTGCTGCTCCATGCTTTTGGCTATGCTGCCGGGCCA 60
Db 1 ATGGGCGCTCGGCTATTTTGTGCTGCTCCATGCTTTTGGCTATGCTGCCGGGCCA 60
QY 61 CGGCGCGGTGACGGGTGATTTCTGAGCGCTGCTGGCGCGGCGGCGGTCCGCGGTGGT 120
Db 61 CGGCGCGGTGACGGGTGATTTCTGAGCGCTGCTGGCGCGGCGGCGGTCCGCGGTGGT 120
QY 121 TTCGCGGTGACGGGTGATTTCTGAGCGCTGCTGGCGCGGCGGCGGTCCGCGGTGGT 180
Db 121 TTCGCGGTGACGGGTGATTTCTGAGCGCTGCTGGCGCGGCGGCGGTCCGCGGTGGT 180
QY 181 CCCTTCGCGCGGTGACGGGTGATTTCTGAGCGCTGCTGGCGCGGCGGCGGTCCGCGGTGGT 240
Db 181 CCCTTCGCGCGGTGACGGGTGATTTCTGAGCGCTGCTGGCGCGGCGGCGGTCCGCGGTGGT 240
QY 241 CGACCACTCGGCTCGGCTGCGGTGACCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 300
Db 241 CGACCACTCGGCTCGGCTGCGGTGACCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 300
QY 301 AGACCTTACACAGCTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 360
Db 301 AGACCTTACACAGCTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 360
QY 361 CCAGTGCCTGATGTCGACCTCCGCGGCGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 420
Db 361 CCAGTGCCTGATGTCGACCTCCGCGGCGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 420
QY 421 TCTCCGCTTTACCTCTTCCGCTGGGCGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 480
Db 421 TCTCCGCTTTACCTCTTCCGCTGGGCGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 480
QY 481 AGTCCGCTTTTACCTCTTCCGCTGGGCGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 540
Db 481 AGTCCGCTTTTACCTCTTCCGCTGGGCGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 540
QY 541 AATTATGCCAGTACCGGCTTCCGCTGGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 600
Db 541 AATTATGCCAGTACCGGCTTCCGCTGGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 600
QY 601 GCTGTCGGGCTTACCGCTTCCGCTGGGCGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 660
Db 601 GCTGTCGGGCTTACCGCTTCCGCTGGGCGGCGGCGGCGGCGGCGGCGGTCCGCGGTGGT 660
QY 661 TCGGTTGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
Db 661 TCGGTTGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720

QY 721 GCTCTGAGCTTGTGATCCAAAGTGAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
Db 721 GCTCTGAGCTTGTGATCCAAAGTGAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
Db 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
QY 841 CATGGCTACTCGTAATTCCTATACTATAACCCCTATACCGGTGCCCTGGGCTGTTG 900
Db 841 CATGGCTACTCGTAATTCCTATACTATAACCCCTATACCGGTGCCCTGGGCTGTTG 900
QY 901 GACTTTGCCCTTGGCTGAGCTTGCACACCTTACCCCGGTAAACCAACCAATACCGGGTC 960
Db 901 GACTTTGCCCTTGGCTGAGCTTGCACACCTTACCCCGGTAAACCAACCAATACCGGGTC 960
QY 961 TCCGCTTATTCAGCACCTGCTCGCCACCGCTTTCGTCGGGTGGGACGGGACTGCCAG 1020
Db 961 TCCGCTTATTCAGCACCTGCTCGCCACCGCTTTCGTCGGGTGGGACGGGACTGCCAG 1020
QY 1021 CTCACCAACCGCTGCTACCCGCTTATGAAGGACCTCTATTTACTAGTACTAATGTT 1080
Db 1021 CTCACCAACCGCTGCTACCCGCTTATGAAGGACCTCTATTTACTAGTACTAATGTT 1080
QY 1081 GTCGCTGAGATCGGCGCGGATAGCCCTCAACCTGCTTCAACCTGCTGACACTCTGCTT 1140
Db 1081 GTCGCTGAGATCGGCGCGGATAGCCCTCAACCTGCTTCAACCTGCTGACACTCTGCTT 1140
QY 1141 GCGCGCTGCGCACAGAAATTTGCTGGGTGGGCGAGCTTCTACTCCCGTCC 1200
Db 1141 GCGCGCTGCGCACAGAAATTTGCTGGGTGGGCGAGCTTCTACTCCCGTCC 1200
QY 1201 GTTGTCTAGCAATGCGGAGCGGCTGTAAAGTGTATACATCTGTAGAGATGCTCAG 1260
Db 1201 GTTGTCTAGCAATGCGGAGCGGCTGTAAAGTGTATACATCTGTAGAGATGCTCAG 1260
QY 1261 CAGGATTAAGGATTAAGCAATCCGATGACATCGGAGAACTCTCGTGTGTTAT 1320
Db 1261 CAGGATTAAGGATTAAGCAATCCGATGACATCGGAGAACTCTCGTGTGTTAT 1320
QY 1321 CAGGATTAAGCAATCAACATGAACATCGGCGAGCGCTTCTCCAGCCCATCGCG 1380
Db 1321 CAGGATTAAGCAATCAACATGAACATCGGCGAGCGCTTCTCCAGCCCATCGCG 1380
QY 1381 CTTTTCTGCTGCTGAGCTAATGATGCTTGGCTCTCTCTACCGCTGCGAGTAT 1440
Db 1381 CTTTTCTGCTGCTGAGCTAATGATGCTTGGCTCTCTCTACCGCTGCGAGTAT 1440
QY 1441 GACCACTCCACTTATGGCTTTCGACTGGCCAGATTTATGTTTCTGACTCTGTGACTTG 1500
Db 1441 GACCACTCCACTTATGGCTTTCGACTGGCCAGATTTATGTTTCTGACTCTGTGACTTG 1500
QY 1501 GTTAATGTTGCAACCGGCGGCGAGCGGCTTGGCGGCTCGCTCGATTTGGAACCAAGTCA 1560
Db 1501 GTTAATGTTGCAACCGGCGGCGAGCGGCTTGGCGGCTCGCTCGATTTGGAACCAAGTCA 1560
QY 1561 CTTGAGCTGCGGCGCTCTCCAGCACTCCAGAGTACTCGAAGACCTTCTTTGCTGCGG 1620
Db 1561 CTTGAGCTGCGGCGCTCTCCAGCACTCCAGAGTACTCGAAGACCTTCTTTGCTGCGG 1620
QY 1621 CTCGCGGTAAAGCTCTCTTTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
Db 1621 CTCGCGGTAAAGCTCTCTTTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
QY 1681 TATAACACCACTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
Db 1681 TATAACACCACTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
QY 1741 ATTTCACCTTACACCACTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
Db 1741 ATTTCACCTTACACCACTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
QY 1801 TTAGCCCCCACTCTGCGCTAGCAATTTGTTGAGGATACCTTGGACTACCTTCCCGCGCC 1860

```
Db 1801 TTAGCCCCCACTCTGCGTAGCATTTGCTTGAGATACCTTGACTACCTGCGCGCC 1860
Qy 1861 CATACTTTTGTGATGATTTTCGCCAGAGTGCGGCCCTTGCGCTTCAGGGCTGCGCTTC 1920
Db 1861 CATACTTTTGTGATGATTTTCGCCAGAGTGCGGCCCTTGCGCTTCAGGGCTGCGCTTC 1920
Qy 1921 CAGTCTACTGCTGAGCTTACGCGCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
Db 1921 CAGTCTACTGCTGAGCTTACGCGCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
Qy 1981 TAG 1983
Db 1981 TAG 1983

RESULT 3
ARI170404
LOCUS ARI170404 2049 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 1 from patent US 6291641.
ACCESSION ARI170404
VERSION ARI170404.1 GI:17908363
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 2049)
AUTHORS Fuerst,R.K., Mcatee,C.,Patrick., Yarbough,P.O. and Zhang,Y.-F.
TITLE Hepatitis E virus antigens and uses therefor
JOURNAL Patent: US 6291641-A 1 18-SEP-2001;
FEATURES
source
1..2049
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 1983; DB 6; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983, Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGCCCTCGGCCATTTTGTGCTGCTCCTCATGTTTTTGTGCTATGCTGCGCGCCA 60
Db 1 ATGCGCCCTCGGCCATTTTGTGCTGCTCCTCATGTTTTTGTGCTATGCTGCGCGCCA 60
Qy 61 CCGCCCGCTCAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120
Db 61 CCGCCCGCTCAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120
Qy 121 TTCTGGGGTGACCGGGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
Db 121 TTCTGGGGTGACCGGGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
Qy 181 CCCTTCGCCGCCGATGTCACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 CCCTTCGCCGCCGATGTCACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Qy 241 CGACCACTCGCTCGCTGCGCGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 300
Db 241 CGACCACTCGCTCGCTGCGCGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 300
Qy 301 AGACCTACCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 AGACCTACCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy 361 CCAGTGCTGATGTCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Db 361 CCAGTGCTGATGTCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Qy 421 TCTCCCTTACTCTTCTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 421 TCTCCCTTACTCTTCTCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Qy 481 AGTCCGCTTTTACCCCTTCAGGACGGCAACCAATACCCATATAATGGCCACGGAAGCTTCT 540
```

```
Db 481 AGTCCGCTTTTACCCCTTCAGGACGGCAACCAATACCCATATAATGGCCACGGAAGCTTCT 540
Qy 541 AATATATCCAGATACCGGGTGGCGGTGCACAAATCCGTTACCGCCCGCTGGTCCCAAT 600
Db 541 AATATATCCAGATACCGGGTGGCGGTGCACAAATCCGTTACCGCCCGCTGGTCCCAAT 600
Qy 601 GCTGTGCGCGGTACGCGCATCTCCATCTCTGCGCCACAGACCAACACCCCGCGCG 660
Db 601 GCTGTGCGCGGTACGCGCATCTCCATCTCTCTGCGCCACAGACCAACACCCCGCGCG 660
Qy 661 TCCGTTGATATGAATTCATAAACCCTCGACGGATGTTGCTATTTTGTGTCAGCCCGGATA 720
Db 661 TCCGTTGATATGAATTCATAAACCCTCGACGGATGTTGCTATTTTGTGTCAGCCCGGATA 720
Qy 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAAACAAAGGTGCGCTCC 780
Db 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAAACAAAGGTGCGCTCC 780
Qy 781 GTGAGACCTCTGCGGTGGCTGAGGAGGAGGCTACCTCTGCTCTGTTGTTGCTTTGCATA 840
Db 781 GTGAGACCTCTGCGGTGGCTGAGGAGGAGGCTACCTCTGCTCTGTTGTTGCTTTGCATA 840
Qy 841 CATGGCTCACTCGTAAATTCCTATATCTATATACACCTTATACCGGTGCCCTCGGCTGTG 900
Db 841 CATGGCTCACTCGTAAATTCCTATATCTATATACACCTTATACCGGTGCCCTCGGCTGTG 900
Qy 901 GACTTTCCTCTGAGCTTGTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACCGGGTC 960
Db 901 GACTTTCCTCTGAGCTTGTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACCGGGTC 960
Qy 961 TCCGTTTATTCAGACCTGCTCGCCACCGCTTCTGCTGCGGTGCGGACGCGACTGCCGAG 1020
Db 961 TCCGTTTATTCAGACCTGCTCGCCACCGCTTCTGCTGCGGTGCGGACGCGACTGCCGAG 1020
Qy 1021 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
Db 1021 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
Qy 1081 GTGCGTGAGATCGCGCGCGGATAGCCCTCACCTGTTTCAACCTTGTGACACTCTGCTT 1140
Db 1081 GTGCGTGAGATCGCGCGCGGATAGCCCTCACCTGTTTCAACCTTGTGACACTCTGCTT 1140
Qy 1141 GCGCGCTGCGCGACAGAAATGATTTCGTCGCTGCTGCGCGCGCTGCTTCTACTCCCGTCCC 1200
Db 1141 GCGCGCTGCGCGACAGAAATGATTTCGTCGCTGCTGCGCGCGCTGCTTCTACTCCCGTCCC 1200
Qy 1201 GTTGTCTCAGCCAAATGGCGAGCGACTGTTAAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Db 1201 GTTGTCTCAGCCAAATGGCGAGCGACTGTTAAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Qy 1261 CAGGATGAAGGTTATGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGTTATT 1320
Db 1261 CAGGATGAAGGTTATGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGTTATT 1320
Qy 1321 CAGGATTTATGATTAACCAACATGAACAGATCGCGCGCGCTTCTCCAGCCCGCATCGCGC 1380
Db 1321 CAGGATTTATGATTAACCAACATGAACAGATCGCGCGCGCTTCTCCAGCCCGCATCGCGC 1380
Qy 1381 CTTTCTCTGTCCTTCGAGCTAATGATGCTGCTTGGCTCTCTCTCACCCTGCGCGAGTAT 1440
Db 1381 CTTTCTCTGTCCTTCGAGCTAATGATGCTGCTTGGCTCTCTCTCACCCTGCGCGAGTAT 1440
Qy 1441 GACCACTGCTGCTTTCGAGCTAATGATGCTGCTTGGCTCTCTCTCACCCTGCGCGAGTAT 1500
Db 1441 GACCACTGCTGCTTTCGAGCTAATGATGCTGCTTGGCTCTCTCTCACCCTGCGCGAGTAT 1500
Qy 1501 GTTAAATGTCGACCGCGCGCGCGCTTGGCGCGCTCGCTCGATTCGATTCGACCAAGGTGACA 1560
Db 1501 GTTAAATGTCGACCGCGCGCGCGCTTGGCGCGCTCGCTCGATTCGATTCGACCAAGGTGACA 1560
Qy 1561 CTTGACCGGTGCGCCCTCTCTCAACATCCAGAGTACTCGAAGACCTTCTTTGTCGCGCG 1620
```

Db 1561 CTTGACGGTGGCCCTCTCCACCATCCAGAGTACTCGAAGACCTCTTTTGTCTCGCG 1620
Qy 1621 CTCGCGGTAGCTCTCTTTCTGGAGGAGGAGCAAACTAAAGCCGGGTACCTTTATAT 1680
Db 1621 CTCGCGGTAGCTCTCTTTCTGGAGGAGGAGCAAACTAAAGCCGGGTACCTTTATAT 1680
Qy 1681 TATAACACACCTGTAGGAGCAAACTGTCTGTGAGAAATGCCGCGGACCGGGTGCCT 1740
Db 1681 TATAACACACCTGTAGGAGCAAACTGTCTGTGAGAAATGCCGCGGACCGGGTGCCT 1740
Qy 1741 ATTTCCACTTACACCACTAGCCTGGGTGCTGCCGCTCTCCATTTCTGCGGTTCGCGTT 1800
Db 1741 ATTTCCACTTACACCACTAGCCTGGGTGCTGCCGCTCTCCATTTCTGCGGTTCGCGTT 1800
Qy 1801 TTAGCGCCCACTCTGCGGTAGCATGCTTGAGGATACCTTGGACTACCTGCCCGCGCC 1860
Db 1801 TTAGCGCCCACTCTGCGGTAGCATGCTTGAGGATACCTTGGACTACCTGCCCGCGCC 1860
Qy 1861 CATACCTTTGATGATTTCTGCCAGAGTGCGCCCTTTGGCTTTCAGGGCTGCGCTTTC 1920
Db 1861 CATACCTTTGATGATTTCTGCCAGAGTGCGCCCTTTGGCTTTCAGGGCTGCGCTTTC 1920
Qy 1921 CAGTCTACTGCTGAGCTTCCAGCGCTTAAAGATCAAGGTGGTAAACTCGGAGTTG 1980
Db 1921 CAGTCTACTGCTGAGCTTCCAGCGCTTAAAGATCAAGGTGGTAAACTCGGAGTTG 1980
Qy 1981 TAG 1983
Db 1981 TAG 1983

RESULT 4
AR232465
LOCUS AR232465 2049 bp mRNA linear PAT 20-DEC-2002
DEFINITION Sequence 3 from patent US 6455492.
ACCESSION AR232465
VERSION AR232465.1 GI:27274534
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2049)
AUTHORS Reyes, G.R., Bradley, D.W., Twu, J.-S., Purdy, M.A., Tam, A.W.,
Krawczynski, K.Z. and Yarbrough, P.O.
TITLE Hepatitis B virus vaccine and method
JOURNAL Patent: US 6455492-A 3 24-SEP-2002;
FEATURES
source Location/Qualifiers
1..2049
/organism="unknown"
/mol_type="mRNA"

ORIGIN
Query Match 100.0%; Score 1983; DB 6; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCCCGGCA 60
Db 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCCCGGCA 60
Qy 61 CCGCCCGGTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 61 CCGCCCGGTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 121 TTCTGGGTGACCGGCTGATTTCTCAGCCCTTCCGAAATCCCTTATATTCATCAACCAAC 180
Db 121 TTCTGGGTGACCGGCTGATTTCTCAGCCCTTCCGAAATCCCTTATATTCATCAACCAAC 180
Qy 181 CCCTTCCCGCCGATGCTACCGCTGCGGCGGGGCTGGACCTGCTTTCGCAACCCGCC 240
Db 181 CCCTTCCCGCCGATGCTACCGCTGCGGCGGGGCTGGACCTGCTTTCGCAACCCGCC 240
Qy 241 CGACCACTCGGCTCGCTTGGCGTGACAGGCCCGCGCCCGCTTGCCTCAGCTCGT 300

Db 241 CGACCACTCGGCTCGCTTGGCGTGACAGGCCCGCGCGCTTGCCTCAGCTCGT 300
Qy 301 AGACCTACACAGCTGGGGCGCGCGCTAACCGCGGTGCTCCGCGCCATGACACCCCG 360
Db 301 AGACCTACACAGCTGGGGCGCGCGCTAACCGCGGTGCTCCGCGCCATGACACCCCG 360
Qy 361 CCAGTGCCTGATGTCGACTCCCGCGGCGCACTCTTTGGCGCGGAGTATTAACCTATCAACA 420
Db 361 CCAGTGCCTGATGTCGACTCCCGCGGCGCACTCTTTGGCGCGGAGTATTAACCTATCAACA 420
Qy 421 TCTCCCTCTTACCTCTTCCGTGCGCAACCGGCACTAACTGTTCTTTATGCGCCCTCTT 480
Db 421 TCTCCCTCTTACCTCTTCCGTGCGCAACCGGCACTAACTGTTCTTTATGCGCCCTCTT 480
Qy 481 AGTCCGCTTTTACCTCTTCCGTGCGCAACCGGCACTAACTGTTCTTTATGCGCCCTCTT 540
Db 481 AGTCCGCTTTTACCTCTTCCGTGCGCAACCGGCACTAACTGTTCTTTATGCGCCCTCTT 540
Qy 541 AATTATGCCAGTACCGGCTTCCCGGTGCGCAACCTGTTACCGCGCTGCTGCCCAAT 600
Db 541 AATTATGCCAGTACCGGCTTCCCGGTGCGCAACCTGTTACCGCGCTGCTGCCCAAT 600
Qy 601 GCTGTGCGGGTTACCGCATCTCCATCTCATTTCTGCGCACAGACCAACACCCCGGAG 660
Db 601 GCTGTGCGGGTTACCGCATCTCCATCTCATTTCTGCGCACAGACCAACACCCCGGAG 660
Qy 661 TCCGTTGATATGAATCAATTAACCTCGAGGATGTTTCTGTTATTTAGTCCAGCCCGGCA 720
Db 661 TCCGTTGATATGAATCAATTAACCTCGAGGATGTTTCTGTTATTTAGTCCAGCCCGGCA 720
Qy 721 GCCTCTGAGCTTGTGATCCCAAGTGCAGCTTACCTATCGTAACCAAGCTGGCGCTCC 780
Db 721 GCCTCTGAGCTTGTGATCCCAAGTGCAGCTTACCTATCGTAACCAAGCTGGCGCTCC 780
Qy 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTA CTTCTGCTTGTGTTATGTTGCA 840
Db 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTA CTTCTGCTTGTGTTATGTTGCA 840
Qy 841 CATGGCTCACTCGTAAATTCCTATATCACTATACCTATACCGCTTACCGGCTGGCTG 900
Db 841 CATGGCTCACTCGTAAATTCCTATATCACTATACCGCTTACCGGCTGGCTG 900
Qy 901 GACTTTGCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 960
Db 901 GACTTTGCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 960
Qy 961 TCCGCTTATTCAGCACTGCTCGCCACCGCTTCTGTCGGGTGCGGACGGGACTGCGGAG 1020
Db 961 TCCGCTTATTCAGCACTGCTCGCCACCGCTTCTGTCGGGTGCGGACGGGACTGCGGAG 1020
Qy 1021 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTCTATTTACTAGTACTATGCT 1080
Db 1021 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTCTATTTACTAGTACTATGCT 1080
Qy 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACTGTTTCAACCTGCTGACACTCTGCTT 1140
Db 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACTGTTTCAACCTGCTGACACTCTGCTT 1140
Qy 1141 GCGCGCTGCGCGACAGAAATTTGATTTCTGCGGTGGTGGCGAGCTGTTCTTCTCCCGTCC 1200
Db 1141 GCGCGCTGCGCGACAGAAATTTGATTTCTGCGGTGGTGGCGAGCTGTTCTTCTCCCGTCC 1200
Qy 1201 GTTGTCTAGCAATGCGGAGCGGACTGTTAGTTGTTATATCTGTAGAGATGCTCAG 1260
Db 1201 GTTGTCTAGCAATGCGGAGCGGACTGTTAGTTGTTATATCTGTAGAGATGCTCAG 1260
Qy 1261 CAGGATAGGGTATTCGAATCCCGATGACTTGGACCTCGGAGATCTCGTGGTTAT 1320
Db 1261 CAGGATAGGGTATTCGAATCCCGATGACTTGGACCTCGGAGATCTCGTGGTTAT 1320
Qy 1321 CAGGATATGATTAACCAATGAATGAATGATGATGATGATGATGATGATGATGATGAT 1380

Db 1126 GTGCTGAGATCGGCGCGGATAGCCCTCACCTGTTCACCTTGTGTGACACTCTGCTT 1185
Qy 1141 GCGGCTCCGACAGAAATGATTTCGTGGCTGGTGGCCAGCTGTCTACTCCGCTCC 1200
Db 1186 GCGGCTCCGACAGAAATGATTTCGTGGCTGGTGGCCAGCTGTCTACTCCGCTCC 1245
Qy 1201 GTTGTCTCAGCCAAATGCGGAGCGACTGTTAAAGTTGTATACATCTGTAGAAATGCTCAG 1260
Db 1246 GTTGTCTCAGCCAAATGCGGAGCGACTGTTAAAGTTGTATACATCTGTAGAAATGCTCAG 1305
Qy 1261 CAGGATAAGGTAATTCGAATCCCGATGACATTCGAGAAATCTCGTGTGTTATT 1320
Db 1306 CAGGATAAGGTAATTCGAATCCCGATGACATTCGAGAAATCTCGTGTGTTATT 1365
Qy 1321 CAGGATTATGATACCAATGACATGACAGATCGCGGAGCGCTTCCAGCCCATCGCG 1380
Db 1366 CAGGATTATGATACCAATGACATGACAGATCGCGGAGCGCTTCCAGCCCATCGCG 1425
Qy 1381 CTTTCTCTGCTCCTCGAGCTAATGATGTGCTTTTGGCTCTCTCACCCTGCGAGTAT 1440
Db 1426 CTTTCTCTGCTCCTCGAGCTAATGATGTGCTTTTGGCTCTCTCACCCTGCGAGTAT 1485
Qy 1441 GACCACTGCACTATGCTCTTCCAGTCCAGTTCATGCTGCTGCTGCTGCTGCTGCTG 1500
Db 1486 GACCACTGCACTATGCTCTTCCAGTCCAGTTCATGCTGCTGCTGCTGCTGCTGCTG 1545
Qy 1501 GTTAATGTTGGACCGCGCGAGCGCTTGGCCGGTGGCTGCTGCTGCTGCTGCTGCTG 1560
Db 1546 GTTAATGTTGGACCGCGCGAGCGCTTGGCCGGTGGCTGCTGCTGCTGCTGCTGCTGCTG 1605
Qy 1561 CTTGACGCTGCGCCCTCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTTGTCTGCG 1620
Db 1606 CTTGACGCTGCGCCCTCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTTGTCTGCG 1665
Qy 1621 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
Db 1666 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1725
Qy 1681 TATAACACCACTGCTAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1726 TATAACACCACTGCTAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1785
Qy 1741 ATTTCACCTTACACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1786 ATTTCACCTTACACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1845
Qy 1801 TTAGCCCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1846 TTAGCCCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1905
Qy 1861 CATACCTTTGATGATTTCTGCCAGAGTGGCCGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1920
Db 1906 CATACCTTTGATGATTTCTGCCAGAGTGGCCGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1965
Qy 1921 CAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Db 1966 CAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2025
Qy 1981 TAG 1983
Db 2026 TAG 2028

RESULT 6
AR232463
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AR232463
Sequence 1 from patent US 6455492.
AR232463
AR232463.1 GI:27274532
Unknown.
Unknown.
Unclassified.

2094 bp
linear
PAT 20-DEC-2002

REFERENCE 1 (bases 1 to 2094)
AUTHORS Reyes, G.R., Bradley, D.W., Tzu, J.-S., Purdy, M.A., Tam, A.W.,
Krawczynski, K.Z. and Yarbough, P.O.
TITLE Hepatitis E virus vaccine and method
JOURNAL Patent: US 645492-A 1 24-SEP-2002;
FEATURES Location/Qualifiers
source 1..2094
/organism="unknown"
/mol_type="mRNA"
ORIGIN
Query Match 100.0%; Score 1983; DB 6; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTTTGGCTATGCTGCGCGGCA 60
Db 46 ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTTTGGCTATGCTGCGCGGCA 105
Qy 61 CGCCCGGTCAGCGCTGCGCGCGCTGCTGCGCGCGGAGCGGCTTCCGCGGTGT 120
Db 106 CGCCCGGTCAGCGCTGCGCGCGCTGCTGCGCGCGGAGCGGCTTCCGCGGTGT 155
Qy 121 TTCTGGGCTGACCGGCTTGAATCTCAGCCCTTCCGAATCCCTATATTCATCAACCAAC 180
Db 166 TTCTGGGCTGACCGGCTTGAATCTCAGCCCTTCCGAATCCCTATATTCATCAACCAAC 225
Qy 181 CCCTTCGCGCGATGTCACCGCTGCGCGCGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 226 CCCTTCGCGCGATGTCACCGCTGCGCGCGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 285
Qy 241 CGACCACTCGGCTCGGCTTGGGTGACAGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCT 300
Db 286 CGACCACTCGGCTCGGCTTGGGTGACAGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCT 345
Qy 301 AGACCTTACCACTGCGCGCGCTGCGCGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCT 360
Db 346 AGACCTTACCACTGCGCGCGCTGCGCGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCT 405
Qy 361 CGAGTCTGATGTCAGCTCCGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCTGCTGCGCGCTGCT 420
Db 406 CGAGTCTGATGTCAGCTCCGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCTGCTGCGCGCTGCT 465
Qy 421 TCTCCCTTTACCTTTCGCTGCGCGCGCTGCGCGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTT 480
Db 466 TCTCCCTTTACCTTTCGCTGCGCGCGCTGCGCGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTT 525
Qy 481 AGTCCGCTTTTACCTTTCGCTGCGCGCGCTGCGCGCTGCGCGCTGCTGCGCGCTGCTGCGCGCT 540
Db 526 AGTCCGCTTTTACCTTTCGCTGCGCGCGCTGCGCGCTGCGCGCTGCTGCGCGCTGCTGCGCGCT 585
Qy 541 AATTATGCCAGTACCGGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCT 600
Db 586 AATTATGCCAGTACCGGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCT 645
Qy 601 GCTGTCGCGGTTTACCGCTTCTCATCTCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCT 660
Db 646 GCTGTCGCGGTTTACCGCTTCTCATCTCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCT 705
Qy 661 TCCGTTGATATGAATTCATTAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 720
Db 706 TCCGTTGATATGAATTCATTAACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 765
Qy 721 GCTCTGAGCTTGTGATTCCTCAAGTACGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 780
Db 766 GCTCTGAGCTTGTGATTCCTCAAGTACGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 825
Qy 781 GTCGAGACCTCTGGGTGCTGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 826 GTCGAGACCTCTGGGTGCTGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 885
Qy 841 CATGGCTCACTCGTAAATTCCTATATACTATACCTTACCTTACCTTACCTTACCTTACCTTACCT 900

```

886 CATGGCTCACTCGTAAATTCCTACTATAATACACCCCTATACCGGTGCCCTCGGGCTGTG 945
901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTACACCAATACCGGGTC 960
946 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTACACCAATACCGGGTC 1005
961 TCCGGTTATTCCAGCACTGCTCGGCACCGCTTCGTGCGGTGGAGCGGACTGCGGAG 1020
1006 TCCGGTTATTCCAGCACTGCTCGGCACCGCTTCGTGCGGTGGAGCGGACTGCGGAG 1065
1021 CTCACCAACCGGTGCTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGT 1080
1066 CTCACCAACCGGTGCTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGT 1125
1081 GTGGTGAGATGCGCGGGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
1126 GTGGTGAGATGCGCGGGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1185
1141 GCGGGCTGCGGACAGAAATTGATTTGCTGGCTGGTGGCCAGCTGTTCTACTCCCGTCCC 1200
1186 GCGGGCTGCGGACAGAAATTGATTTGCTGGCTGGTGGCCAGCTGTTCTACTCCCGTCCC 1245
1201 GTTGTCTCAGCAATGCGGACCGCACTGTTAAAGTTGTATATCATCTGAGAGAATGCTCAG 1260
1246 GTTGTCTCAGCAATGCGGACCGCACTGTTAAAGTTGTATATCATCTGAGAGAATGCTCAG 1305
1261 CAGGATAAGGTTATGCAATCCCGCAGTACATTCACCTCGGAGAACTCTGTTGTTATT 1320
1306 CAGGATAAGGTTATGCAATCCCGCAGTACATTCACCTCGGAGAACTCTGTTGTTATT 1365
1321 CAGGATTATGATAACCAACATGAACATCGGCGGACGCTTCTCAGGCCCAATCGCGC 1380
1366 CAGGATTATGATAACCAACATGAACATCGGCGGACGCTTCTCAGGCCCAATCGCGC 1425
1381 CTTTCTCTGTCCTTCGAGCTAATGATGTGCTTTGGTCTCTCTCAACGCTGCGGAGTAT 1440
1426 CTTTCTCTGTCCTTCGAGCTAATGATGTGCTTTGGTCTCTCTCAACGCTGCGGAGTAT 1485
1441 GACAGTCCACTTATGCTCTTCACTGGCCAGCTTTATGTTCTGACTCTGTGACCTTG 1500
1486 GACAGTCCACTTATGCTCTTCACTGGCCAGCTTTATGTTCTGACTCTGTGACCTTG 1545
1501 GTTAAATGTTGGAACCGGCGGACGCTTTCGCGGTGCTCGATTGGACCAAGGTACA 1560
1546 GTTAAATGTTGGAACCGGCGGACGCTTTCGCGGTGCTCGATTGGACCAAGGTACA 1605
1561 CTTGACGGTCCGCCCTCTCCACCATCAGCAGTACTCGAAGACTTCTTTGTCTCGCGG 1620
1606 CTTGACGGTCCGCCCTCTCCACCATCAGCAGTACTCGAAGACTTCTTTGTCTCGCGG 1665
1621 CTCGCGGTAGCTCTCTTTCTGGGAGCGGACCACTAAGCGGGTACCCCTATAAT 1680
1666 CTCGCGGTAGCTCTCTTTCTGGGAGCGGACCACTAAGCGGGTACCCCTATAAT 1725
1681 TATAACCACTAGCTAGCGACCACTGCTTGTGAGAAATGCGCGGACCGGCTCGCT 1740
1726 TATAACCACTAGCTAGCGACCACTGCTTGTGAGAAATGCGCGGACCGGCTCGCT 1785
1741 ATTTCCACTTACACCACTAGCTAGCGACCACTGCTTGTGAGAAATGCGCGGCTCGCT 1800
1786 ATTTCCACTTACACCACTAGCTAGCGACCACTGCTTGTGAGAAATGCGCGGCTCGCT 1845
1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTGCAGATACCTTGGACTACCTGCGCGGCG 1860
1846 TTAGCCCCCACTCTGCGCTAGCAATGCTTGCAGATACCTTGGACTACCTTGC CGCGGCG 1905
1861 CATACTTTTGTATGATTTCTGCGGAGTGCCTGCTCCATTTCTGCGGTTGCGGTT 1920
1906 CATACTTTTGTATGATTTCTGCGGAGTGCCTGCTCCATTTCTGCGGTTGCGGTT 1965
1921 CAGTCTACTGTCGCTGAGCTTACGCGCTTAAAGTGAAGGTGGGTAAATCTCGGGAGTTG 1980
1966 CAGTCTACTGTCGCTGAGCTTACGCGCTTAAAGTGAAGGTGGGTAAATCTCGGGAGTTG 2025

RESULT 7
LOCUS ARL150915 7195 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 6 from patent US 6229005.
ACCESSION ARL150915
VERSION ARL150915.1 GI:15115506
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7195)
AUTHORS Reyes,G.R., Yarbough,P.O., Bradley,D.W., Krawczynski,K.Z., Tam,A.W.
and Fry,K.E.
TITLE DNA sequences of enterically transmitted non-A/non-B hepatitis
viral agent
JOURNAL Patent: US 6229005-A 6 08-MAY-2001;
FEATURES
source Location/Qualifiers
1..7195
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 1983; DB 6; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTTCCTATGCTGCGCGCGCA 60
Db 5147 ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTTCCTATGCTGCGCGCGCA 5206
Qy 61 CCGCGCCGTCAGCGCTGCGCGCGCGTCGTGGGGCGGCGAGCGCGGTTCGCGGGTGGT 120
Db 5207 CCGCGCCGTCAGCGCTGCGCGCGCGTCGTGGGGCGGCGAGCGCGGTTCGCGGGTGGT 5266
Qy 121 TTTGCGGGTGACCGGGTTGATTCTCAGCCCTTTCGCAATTCCTATATTCATCCAAACAC 180
Db 5267 TTTGCGGGTGACCGGGTTGATTCTCAGCCCTTTCGCAATTCCTATATTCATCCAAACAC 5326
Qy 181 CCTTCCGCGCCGATGTCACCGCTGCGCGCGGCTGAGCTCTGTTTCGCAACCGCGC 240
Db 5327 CCTTCCGCGCCGATGTCACCGCTGCGCGCGGCTGAGCTCTGTTTCGCAACCGCGC 5386
Qy 241 CGACCACTCGGCTCGGCTTGGCGTGACCGCGCGCGCGCGCGCGCTTGCCTCAGTCGT 300
Db 5387 CGACCACTCGGCTCGGCTTGGCGTGACCGCGCGCGCGCGCGCGCTTGCCTCAGTCGT 5446
Qy 301 AGACCTACCAAGTGGGGCGCGCGCTAACCGCGGTGCTCCGGCGCCATGACACCCCG 360
Db 5447 AGACCTACCAAGTGGGGCGCGCGCTAACCGCGGTGCTCCGGCGCCATGACACCCCG 5506
Qy 361 CCAGTGCTGATGTCGACTCCCGCGCGCATCTTGGCGCGGAGTAACTTAACTCAACA 420
Db 5507 CCAGTGCTGATGTCGACTCCCGCGCGCATCTTGGCGCGGAGTAACTTAACTCAACA 5566
Qy 421 TCTCCCTTACCTCTCCGTGGCCACCGGCACTAACCTGGTTCCTTTATGCGCGCCCTCT 480
Db 5567 TCTCCCTTACCTCTCCGTGGCCACCGGCACTAACCTGGTTCCTTTATGCGCGCCCTCT 5626
Qy 481 AGTCGGTTTTTACCCCTTCAGACGGCGCACCAATACCATATATGCGCGGAGCTTCT 540
Db 5627 AGTCGGTTTTTACCCCTTCAGACGGCGCACCAATACCATATATGCGCGGAGCTTCT 5686
Qy 541 AATTATGCCAGTACCGGGTTGCGCGTGCACCAATTCGCTTACCGCGCGCTGGTCCCCAAT 600
Db 5687 AATTATGCCAGTACCGGGTTGCGCGTGCACCAATTCGCTTACCGCGCGCTGGTCCCCAAT 5746
Qy 601 GCTGTCGGCGGTTACGGCAATCTCCATCTCATTTCTTGGGCCACAGACCAACCCCGGAG 660
```



```
Db 5507 CCAAGTGCCTGATGTGCAGTCTCCGCGCGCCATCTTGGCGCGCAGTATAACCTATCAACA 5566
Qy 421 TCTCCCTTACCTCTTCGCGTGGGCAACCGGCACTAAACCTGTGTTCTTTATGCGGCCCTCTT 480
Db 5567 TCTCCCTTACCTCTTCGCGTGGGCAACCGGCACTAAACCTGTGTTCTTTATGCGGCCCTCTT 5626
Qy 481 AGTCCGCTTTACCCCTTCAGGAGCGGCAACCAATACCATATATAATGGCCACGGAAGCTTCT 540
Db 5627 AGTCCGCTTTACCCCTTCAGGAGCGGCAACCAATACCATATATAATGGCCACGGAAGCTTCT 5686
Qy 541 AATTATGCCAGTACCGGGTGGCCGTGGCCACCAATCCGTTACCGCCCGCTGGTCCCAAT 600
Db 5687 AATTATGCCAGTACCGGGTGGCCGTGGCCACCAATCCGTTACCGCCCGCTGGTCCCAAT 5746
Qy 601 GCTGTGGCGGGTACGGCATCTGCATCTCAATCTTCTGTCGACAGACAGACCAACCCGAGC 660
Db 5747 GCTGTGGCGGGTACGGCATCTGCATCTCAATCTTCTGTCGACAGACAGACCAACCCGAGC 5806
Qy 661 TCCGTTGATATGAATCAATAACCTCGACGGATGTTTCGTATTTTAGTCCAGCGCGGCATA 720
Db 5807 TCCGTTGATATGAATCAATAACCTCGACGGATGTTTCGTATTTTAGTCCAGCGCGGCATA 5866
Qy 721 GCCTCTGAGCTTGATGCCAAGTGAGCGCTTACACTATCGTAACCAAGGCTGGCGTCC 780
Db 5867 GCCTCTGAGCTTGATGCCAAGTGAGCGCTTACACTATCGTAACCAAGGCTGGCGTCC 5926
Qy 781 GTCCAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCGTCTGTGTTATGCTTTGCATA 840
Db 5927 GTCCAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCGTCTGTGTTATGCTTTGCATA 5986
Qy 841 CATGGCTCACTCGTAATTCCTTACTATACTAACACCTTATACCGGTGCGCTCGGCTGTTG 900
Db 5987 CATGGCTCACTCGTAATTCCTTACTATACTAACACCTTATACCGGTGCGCTCGGCTGTTG 6046
Qy 901 GACTTTGCCCTTGAGCTTGATTTGCCAACCTTTACCCCGGTAAACCAATACCGCGGTC 960
Db 6047 GACTTTGCCCTTGAGCTTGATTTGCCAACCTTTACCCCGGTAAACCAATACCGCGGTC 6106
Qy 961 TCCCGTATTCAGCACTGCTCGCCACCGCTTGTGCGGGTGGGACGGAGCTGGCGAG 1020
Db 6107 TCCCGTATTCAGCACTGCTCGCCACCGCTTGTGCGGGTGGGACGGAGCTGGCGAG 6166
Qy 1021 CTCACCAACACGGCTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGT 1080
Db 6167 CTCACCAACACGGCTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGT 6226
Qy 1081 GTCGGTGAGATCGGCCCGGGATAGCCCTCACCTGTTCAACCTTGTGACACACTCTGCTT 1140
Db 6227 GTCGGTGAGATCGGCCCGGGATAGCCCTCACCTGTTCAACCTTGTGACACACTCTGCTT 6286
Qy 1141 GCGGGCTGCGGACAGAAATGATTCGTGGCTGTTGGCCAGCTGTTCTACTCCGCTCC 1200
Db 6287 GCGGGCTGCGGACAGAAATGATTCGTGGCTGTTGGCCAGCTGTTCTACTCCGCTCC 6346
Qy 1201 GTTGCTTCAGCCAAATGCGGACCGACAGTGTAAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 6347 GTTGCTTCAGCCAAATGCGGACCGACAGTGTAAAGTTGTATACATCTGTAGAGAATGCTCAG 6406
Qy 1261 CAGGATAAGGGTATTTGAATCCCGCATGACATTTGATGATCTGTAGAGAATGCTCAG 1320
Db 6407 CAGGATAAGGGTATTTGAATCCCGCATGACATTTGATGATCTGTAGAGAATGCTCAG 6466
Qy 1321 CAGGATTTATGATAACCAATGAACAGATCGGCCAGCGCTTCTCCAGCCCATCGGC 1380
Db 6467 CAGGATTTATGATAACCAATGAACAGATCGGCCAGCGCTTCTCCAGCCCATCGGC 6526
Qy 1381 CCTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCAACCGCTGCCGAGTAT 1440
Db 6527 CCTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCAACCGCTGCCGAGTAT 6586
Qy 1441 GACAGTCCACTTATGGCTCTTGA CTGGGCCAGTTATGTTTCTGACTCTGTGACCTTG 1500
```

```
6587 GACCAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 6646
1501 GTTAAATGTTGACCGGCGGAGGGCGTTGCGCGGTGCTCGATTGGACCAAGGTACA 1560
6647 GTTAAATGTTGACCGGCGGAGGGCGTTGCGCGGTGCTCGATTGGACCAAGGTACA 6706
1561 CTTGAGCGTGGCCCTCTCCACCATCCAGCACTGCTCGAAGACCTTCTTTGCTCTCGCG 1620
6707 CTTGAGCGTGGCCCTCTCCACCATCCAGCACTGCTCGAAGACCTTCTTTGCTCTCGCG 6766
1621 CTCGCGGTAAGCTCTCTTTTCTGGAGGCGAGGACAACTAAAGCCGGGTACCTTATAAT 1680
6767 CTCGCGGTAAGCTCTCTTTTCTGGAGGCGAGGACAACTAAAGCCGGGTACCTTATAAT 6826
1681 TATAACACCACTGCTAGCGACCACTGCTTGTGAGAAATCGCGCGGGACACGGGTGCT 1740
6827 TATAACACCACTGCTAGCGACCACTGCTTGTGAGAAATCGCGCGGGACACGGGTGCT 6886
1741 ATTTCACATTACACCACTAGCTGCTGGGTGCTGGTCCCGTCTCCATTTCTCGGTTGCCGTT 1800
6887 ATTTCACATTACACCACTAGCTGCTGGGTGCTGGTCCCGTCTCCATTTCTCGGTTGCCGTT 6946
1801 TTAGCCCCCACTCTCGCTAGCATCTTTAGGATACCTTTGACTACCTGCCCGCGCC 1860
6947 TTAGCCCCCACTCTCGCTAGCATCTTTAGGATACCTTTGACTACCTGCCCGCGCC 7006
1861 CATACTTTGATGATTTGCCCCAGAGTGGCGGCCCTTGGCCCTTACGGGCTGGCTTTC 1920
7007 CATACTTTGATGATTTGCCCCAGAGTGGCGGCCCTTGGCCCTTACGGGCTGGCTTTC 7066
1921 CAGTCTACTGCTCGCTGAGCTTACGCGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
7067 CAGTCTACTGCTCGCTGAGCTTACGCGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 7126
1981 TAG 1983
7127 TAG 7129

RESULT 9
HPESVP
LOCUS
DEFINITION
Hepatitis E structural and nonstructural protein mRNA, complete
cds'
ACCESSION
M73218.1 GI:330023
VERSION
nonstructural protein; viral structural protein.
KEYWORDS
Hepatitis E virus
SOURCE
Hepatitis E virus
ORGANISM
Viruses; ssRNA positive-strand viruses, no DNA stage; Hepatitis
E-like viruses.
REFERENCE
1 (bases 1 to 7207)
AUTHORS
Tam,A.W., Smith,M.M., Guerra,M.E., Huang,C.C., Bradley,D.W.,
Fry,K.E. and Reyes,G.R.
TITLE
Hepatitis E virus (HEV): molecular cloning and sequencing of the
full-length viral genome
JOURNAL
Virology 185 (1), 120-131 (1991)
MEDLINE
92024067
PubMed
1926770
COMMENT
Original
FEATURES
Location/Qualifiers
source
1..7207
/organism="Hepatitis E virus"
/mol_type="genomic RNA"
/db_xref="taxon:12461"
28..5109
/notes="putative"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA45734.1"
/db_xref="GI:330024"
source text: Hepatitis E virus cDNA to genomic RNA.
Location/Qualifiers
1..7207
/organism="Hepatitis E virus"
/mol_type="genomic RNA"
/db_xref="taxon:12461"
28..5109
/notes="putative"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAA45734.1"
/db_xref="GI:330024"
translation="MEAHQFIKAPGTTTAEQAALAAANSALANAVVPRFLSHQQIE
ILINLMQRLVFRPEVFNHPQIRVIHNELEYCRSGRCLEIGHPRINDPNV
```

VHRCFLRPVGRDVRWYAPTGRPAANCRRSALRGLPAADRTYCLDGFSGNFPBAETG
IALYSIHOMSPSDVAEMPRHCWTRLVYALHLPPEVLLPPGTYRTASVLLIHDGRVV
VYIEGDTSGNHDVSNLRSWTRTKYTGCDHDLVIERVRAIGCHVFILLITAAPEBSPM
PIVYPRSTEVTVRSISFGGGTSPSFPSTSCSTGTHLPIERVAPHAIDWRLMFLGAILDQQA
FCCSLMTYLRIGSYKVTYVTVIANEGMNASDALTAIVTAAYLTICHQRYLRTQALIS
KMRRLERHAQKFTIRLYSWIFKXWGOECIFLOQLEFYAOCRRWLSAGFHLDPRLV
PESASLDISGVSYPVGTALQPLQALDLPALAEIVARAGRLTATVKVSDVDIGDETL
LGNKTRFSFUDGAVLENGPERHNLSPDASQTMAGPFLSYTAAASAGLEVRVVA
GLDRAVAPGSPSPAGEVTAFCALYRFNREARQSHLSIGNLWFHPEGLGLTAPF
SLOHVESANPFCESTYLRTWSEVDVASSPARDLGFMSEPSIPSRAAATPLAALP
PPAPDPSPPADLAEPASGATACAPAIHQATARHRLLETFYDPSKVFAGSFLFES
TCTLWLNASVNDPREGGLCHAFYQRYPASFDAAFWMRDGAAYTLTPRPIIHAVAP
DYLRHNKRLEAAVRETCSLRGTAAYPILGTYGVPTGSPSFDWERNHRPGDELXL
PILAAWPEANPREPTLTITIEDVARTANLAJELDSATDVGRACGRVTPGVQYOF
TAGVFGSKRSITQADVVDVVVPTRELNWRNRGFAAFPHTHAARTVQGRVVIDE
APSLPHLLHLMQAAATVHLLGDPNQIPADIDFEHAGIVPAIRPDLPSTVMMHWTRW
PADVCELIRGTPMOTTSVRLSLFWGEPVQKLVFTQAAKPNAGSIVTHEAQA
TYTETIITADARGLIQSSRAHAIVALTRHTEKCVIIDAPGLLREVGISDAIVNFF
LAGELIGHORPSVIRPGNPANDVTLLAAPPSQISAPHOLAEELGHRPVAVALPP
CPLEBQGLLYLQELTTCDSVUTFEITDVIHCEMAAPSORKAVLTLVGRYGRKLY
NASHDVRSLARFIPAIGVQVVTTCLELYELVEAMVEKQDGSVLELDELNCVRSI
TFQDKCNKFTGETIAHKVGCGISAMSKTFCALFGPMFRAIEKAILALLPQGVYF
DAFDOTFSAVAAKASAMFENDEFSDTQNNFSLEICAMEECQPMQWILRLYH
LIRSAWILQAPKESIRGFKWHSBPGTLLWNTVMNAVITHCYDFRDPQVAAFGRDD
STVLCEYRQSPGAALVLAGCGLKLVDPFRPLGLVAGVVVAPGLGALPDVVFAGRLT
EKWNGPBRAPRAQLAVSDFLKLTNVAMQMDVVSVRYGSPGLVHNLIGMLQAVA
DGAHFTESVKPVLDTNLSILCRVE"
5106..5477
/codon_start=1
/product="structural viral protein"
/protein_id="AA45735.1"
/db_xref="GI:330025"
/translation="MNNMSFAAPMGSRPCALGLFCCSCFCCLCPRHPRVSRLLAVV
GGAAPVAVSVGTGLIUSPSQSPFIQPTSPMSPURPGLDLVFNANPDHSAPLGV
TRPSAPPLPHVVDLPQLGPRR"
5147..7129
/codon_start=1
/product="structural viral protein"
/protein_id="AA45736.1"
/db_xref="GI:330026"
/translation="WRPRPILLLLMLPMLPAPPQGPQSPRRRRSGSGSGFGWD
RVDSOPFATPIHTNPPADVPVQAPRPLGSAWRQAPAVASRRRP
TTAGAPLTAAPADPTVPDVSRAILRQYNLSTSPLTSSVATQNLVLVYAAPL
SELLPDGNTNTHIMATESNVAQIRVARATIRPLVNAVGVTAISISFWPQTITI
PISVDMNLSITDRLVQPIASELVIPISERLHYNQWRSVETSGVAREEATSLV
MLCISGLVNSYNTPIGALGLDFALEFNLTPGNTNTRVSRYSSTARHLRRG
ADGTAELETTAATRFMKDLYFTNGVGEIGRIGIALTLFNLADTLGLLPLELISAG
GQLFYSRPVVSANGEPVKLYTSVENAQODKGIATPHDIDLGESRVVIQDYDNQHEQD
RPTSPAPSRPESVLRANDVLSLTAAYDQSTYGSSTGSPVVSVDSTVLNVNITAO
AVARSLDWTKVTLDRPPLSTIOQYSKTFVPLPLRGKLSPEAGTTKAGYPVNYNTAS
DQLLVENAAHGRAVIAISTYTSIGAGVSVISAVAVLAPHASALALLEDTLDYPARHTFD
DFPCPRPLUGLQCAQFQSTVABELQRLMKVKGKREL"

ORIGIN

Query Match 100.0%; Score 1983; DB 14; Length 7207;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGCGCTATTTTGTGCTGCTCCTCATGTTTTTGGCTATGCTGCCCGGCCA 60
DB 5147 ATGCGCCCTCGCGCTATTTTGTGCTGCTCCTCATGTTTTTGGCTATGCTGCCCGGCCA 5206
QY 61 CGCCCGGTCAGCGCTGTCGCCCGCTGTCGTGGCGCGGAGCGCGGTTCGCGCGGTGTT 120
DB 5207 CGCCCGGTCAGCGCTGTCGCCCGCTGTCGTGGCGCGGAGCGCGGTTCGCGCGGTGTT 5266
QY 121 TTCTGGGGTGACCGGGTTCATCTCAGCCTTCGCAATCCCTATATTTCATCAACCAAC 180
DB 5267 TTCTGGGGTGACCGGGTTCATCTCAGCCTTCGCAATCCCTATATTTCATCAACCAAC 5326
QY 181 CCCTTCGCCCGGATGTCACCGCTGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCG 240

781 GTCCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGCTTGTGTTATGCTTTGCATA 840
841 CATGGCTCAGTCGTAATTCCTATACCTATACACCTATACCGGTGCCCTCGGGCTGTTG 900
841 CATGGCTCAGCCGTAATTCCTATACCTATACACCTATACCGGTGCCCTCGGGCTGTTG 900
901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC 960
901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC 960
961 TCCCGTTATTCAGCACTGCTCGCCACCGCCCTTCGTGCGGTGCGGAGGCACTCCCGAG 1020
961 TCCCGTTATTCAGCACTGCTCGCCACCGCCCTTCGTGCGGTGCGGAGGCACTCCCGAG 1020
1021 CTCACACACGCTGCTACCGCTTTATGAAGACCTCTATTTACTAGTACTAATGGT 1080
1021 CTCACACACGCTGCTACCGCTTTATGAAGACCTCTATTTACTAGTACTAATGGT 1080
1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
1141 GSCGCGCTGCCACAGAAATGATTTGTCGGCTGCTGCGGAGCTGTTACTTCCCGTCCC 1200
1141 GSCGCGCTGCCACAGAAATGATTTGTCGGCTGCTGCGGAGCTGTTACTTCCCGTCCC 1200
1201 GTTGTCTAGCAATGCGGAGCGGACTGTTAGTTGTATACATCTGTAGAGAACTCTCAG 1260
1201 GTTGTCTAGCAATGCGGAGCGGACTGTTAGTTGTATACATCTGTAGAGAACTCTCAG 1260
1261 CAGGATAGGATTTGCAATCCGATGACATGACCTCGGAGAACTCTGCTGCTGTTAT 1320
1261 CAGGATAGGATTTGCAATCCGATGACATGACCTCGGAGAACTCTGCTGCTGTTAT 1320
1321 CAGGATTTATGATPACCAACATGAACATGACGCGGAGCTTCTCCAGCCCGATCGCGC 1380
1321 CAGGATTTATGATPACCAACATGAACATGACGCGGAGCTTCTCCAGCCCGATCGCGC 1380
1381 CTTTCTCTGCTCTGAGCTAATGATGCTTTGCTGCTCTCTCACCGCTGCCAGTAT 1440
1381 CTTTCTCTGCTCTGAGCTAATGATGCTTTGCTGCTCTCTCACCGCTGCCAGTAT 1440
1441 GACCACTCAGCTTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
1441 GACCACTCAGCTTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
1501 GTTAATGTTGACCGCGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
1501 GTTAATGTTGACCGCGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
1561 CTTGACGCTGCGCGCTCTGACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCTGCG 1620
1561 CTTGACGCTGCGCGCTCTGACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCTGCG 1620
1621 CTTCCGCGTAACTCTCTTTCTGGGAGGAGGACCACTAAAGCCGGGTACCTTATATAT 1680
1621 CTTCCGCGTAACTCTCTTTCTGGGAGGAGGACCACTAAAGCCGGGTACCTTATATAT 1680
1681 TATAACACCACTGCTAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1681 TATAACACCACTGCTAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1741 ATTTCACCTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1741 ATTTCACCTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1801 TTAGCCCCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1801 TTAGCCCCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1861 CATACCTTTGATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
1861 CATACCTTTGATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920

1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAGATGAAGTGGTAAACTCGGAGTTA 1980
1981 TAG 1983
1981 TAG 1983
RESULT 11
E17109
LOCUS Hepatitis E virus gene. 7194 bp DNA linear PAT 28-JUL-1999
DEFINITION Hepatitis E virus gene.
ACCESSION E17109
VERSION E17109.1 GI:5711792
KEYWORDS JP 1998234383-A/5.
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 7194)
AUTHORS Takeda,N., Ri,T and Miyamura,T.
TITLE HEPATITIS E VIRUS HOLLOW PARTICLE, GENE ENCODING THE SAME AND PREPARATION OF RECOMBINATION VECTOR CONTAINING THE GENE AND UTILIZATION OF THE RECOMBINATION VECTOR
JOURNAL Patent: JP 1998234383-A 5 08-SEP-1998;
COMMENT DENKA SEIKEN CO LTD, KOKURITSU KANSENSHIYOU KENKYUSHO
OS Hepatitis E virus
PN JP 1998234383-A/5
PD 08-SEP-1998
PF 28-FEB-1997 JP 1997062445
PI TAKEDA NAKAZU, RI TENSEI, MIYAMURA TATSUO
PC C12N15/09,C07K14/08,C12N5/10,C12P21/02;
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..7194
FT /organism='Hepatitis E virus'.
FEATURES
source
1..7194
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 98.3%; Score 1949.4; DB 6; Length 7194;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1962; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCGCGCCA 60
DB 5147 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCGCGCCA 5206
QY 61 CGCGCCGCTCAGCGCTCTGCGCCCGCTGCGCGCGCGCGCGGTTCCGCGGTGT 120
DB 5207 CGCGCCGCTCAGCGCTCTGCGCCCGCTGCGCGCGCGCGGTTCCGCGGTGT 5266
QY 121 TTCTGGGTGACCGGGTTGATTTCTAGCCCTTCGCAATCCCTATATTTTCATCAACCAAC 180
DB 5267 TTCTGGGTGACCGGGTTGATTTCTAGCCCTTCGCAATCCCTATATTTTCATCAACCAAC 5326
QY 181 CCCTTCGCCCCGATGTACCGCTCGCGCGGGGTGACCTCTGTTTCGCAACCGCGC 240
DB 5327 CCCTTCGCCCCGATGTACCGCTCGCGCGGGGTGACCTCTGTTTCGCAACCGCGC 5386
QY 241 CGACCACTCGGCTCGGCTTTGGGTGACAGGCGCCAGCGCCCGCTTGCCTCAGCTGT 300
DB 5387 CGACCACTCGGCTCGGCTTTGGGTGACAGGCGCCAGCGCCCGCTTGCCTCAGCTGT 5446
QY 301 AGACCTTACCAAGTGGGCGCGCGCTTAACCGCGGTGCTCCGCGCCCATGACACCCCG 360
DB 5447 AGACCTTACCAAGTGGGCGCGCGCTTAACCGCGGTGCTCCGCGCCCATGACACCCCG 5506

QY 361 CCAGTGCCTGATGTCGACTCCCGGCGCCCATCTTTGGCGGCGAGTATAACCTATCAACA 420
Db |||||
QY 5507 CCAGTGCCTGATGTCGACTCTCGGCGGCCCATCTTTGGCGGCGAGTATAATCTATCAACA 5566
Db |||||
QY 421 TCTCCCTTACCTCTTCCGCGGCGACCGGCACTAACCTGGTCTCTTTATGCGCGCCCTCTT 480
Db |||||
QY 5567 TCTCCCTACCTCTTCCGCGGCGACCGGCACTAACCTGGTCTCTTTATGCGCGCCCTCTT 5626
Db |||||
QY 481 AGTCCGCTTTTACCCCTTACGAGCGGACCAATACCCATATATATGCGCACGGAAGCTTCT 540
Db |||||
QY 5627 AGTCCGCTCTTACCTCTTACGAGCGGACCAATACCCATATATATGCGCACGGAAGCTTCT 5686
Db |||||
QY 541 AATTATGCCAGTACCGGCTTGCCTGCGCACCAATCCGTTTACCGCGCTGGTCCCAAT 600
Db |||||
QY 5687 AATTATGCCAGTACCGGCTTGCCTGCGCACCAATCCGTTTACCGCGCTGGTCCCAAT 5746
Db |||||
QY 601 GCTGTGCGGCTTACGCGCATCTCATCTCATCTGCGCACAGACCAACCAACCGCGAGC 660
Db |||||
QY 5747 GCTGTGCGGCTTATGCGCATCTCATCTCATCTGCGCACAGACCAACCAACCGCGAGC 5806
Db |||||
QY 661 TCCGTTGATGATGAATTCGAATTCGAGCGGCTTACACTATCGTAACCAAGGCTGGCGTCC 780
Db |||||
QY 5867 GCCTCTGAGCTTGATCCCAAGTGAGCGGCTTACACTATCGTAACCAAGGCTGGCGTCC 5926
Db |||||
QY 781 GTCGAGACCTCTGGGCTGCTGAGAGGAGGCTTACCTCTGGTCTTGTATGCTTTGCATA 840
Db |||||
QY 5927 GTCGAGACCTCTGGGCTGCTGAGAGGAGGCTTACCTCTGGTCTTGTATGCTTTGCATA 5986
Db |||||
QY 841 CATGGCTCACTCGTAATTCCTATCTATACACCCCTATACCGTGGCTTGGCTGGCTG 900
Db |||||
QY 5987 CATGGCTCACTCGTAATTCCTATCTATACACCCCTATACCGTGGCTTGGCTGGCTG 6046
Db |||||
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACTTACCCCGGTAACCAATACCGCGGTC 960
Db |||||
QY 6047 GACTTTGCCCTTGAGCTTGAGTTTCGCAACTTACCCCGGTAACCAATACCGCGGTC 6106
Db |||||
QY 961 TCCGTTATTCAGCACTGCTGCCACCGCTTTCGCGGCTGCGGAGCGGAGCTGCCGAG 1020
Db |||||
QY 6107 TCCGTTATTCAGCACTGCTGCCACCGCTTTCGCGGCTGCGGAGCGGAGCTGCCGAG 6166
Db |||||
QY 1021 CTCACCAACGCTGCTACCGCTTATGAGGAGCTCTATTTTACTAGTACTATGTT 1080
Db |||||
QY 6167 CTCACCAACGCTGCTACCGCTTATGAGGAGCTCTATTTTACTAGTACTATGTT 6226
Db |||||
QY 1081 GTCGGTGAGATCGGCGCGGAGTAGCCCTCACCTGTTCAACCTTGTCTGACACTCTGCTT 1140
Db |||||
QY 6227 GTCGGTGAGATCGGCGCGGAGTAGCCCTCACCTGTTTAACTTGTCTGACACTCTGCTT 6286
Db |||||
QY 1141 GCGGCTGCTCCAGCAATGATTTGCTGCGTGGTGGCGAGCTGTTTACTCCGCTCC 1200
Db |||||
QY 6287 GCGGCTGCTCCAGCAATGATTTGCTGCGTGGTGGCGAGCTGTTTACTCCGCTCC 6346
Db |||||
QY 1201 GTTGTCTCAGCAATGGCGAGCGGCTGTTAGTGTGTATACATCTGTAGAGAATGCTAG 1260
Db |||||
QY 6347 GTTGTCTCAGCAATGGCGAGCGGCTGTTAGTGTGTATACATCTGTAGAGAATGCTAG 6406
Db |||||
QY 1261 CAGGATAAGGTTATGCAATCCCGCATGACATTTGACCTCGGAGAATCTCGTGGTATT 1320
Db |||||
QY 6407 CAGGATAAGGTTATGCAATCCCGCATGACATTTGACCTCGGAGAATCTCGTGGTATT 6466
Db |||||
QY 1321 CAGGATATGATAACCAATGACCAAGATCGGCGAGCGCTTCTCAGCCCATCGCGC 1380
Db |||||
QY 6467 CAGGATATGATAACCAATGACCAAGATCGGCGAGCGCTTCTCAGCCCATCGCGC 6526
Db |||||
QY 1381 CTTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT 1440
Db |||||
QY 6527 CTTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT 6586
Db |||||

QY 1441 GACAGTCCACTTATGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db |||||
QY 6587 GACAGTCCACTTATGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 6646
Db |||||
QY 1501 GTTAAATGTTTCGACCGCGCGGCGGCTTTCGCCGCTCGCTCGATTGGACCAAGGTCACA 1560
Db |||||
QY 6647 GTTAAATGTTTCGACCGCGCGGCGGCTTTCGCCGCTCGCTCGATTGGACCAAGGTCACA 6706
Db |||||
QY 1561 CTTGAGCGTGGCGCCCTCTCCACATCCAGAGTACTCGAAGACCTTCTTTGCTCTCCG 1620
Db |||||
QY 6707 CTTGAGCGTGGCGCCCTCTCCACATCCAGAGTACTCGAAGACCTTCTTTGCTCTCCG 6766
Db |||||
QY 1621 CTCGCGGTAAGCTCTCTTTCGAGGAGGAGGACCACTAAAGCCGGTACCTTATAAT 1680
Db |||||
QY 6767 CTCGCGGTAAGCTCTCTTTCGAGGAGGAGGACCACTAAAGCCGGTACCTTATAAT 6826
Db |||||
QY 1681 TATAACACCACTGCTTAGCGACCAACTGTTTGTGAGAAATGCCGCCGCGGACCGGCTCGCT 1740
Db |||||
QY 6827 TATAACACCACTGCTTAGCGACCAACTGTTTGTGAGAAATGCCGCCGCGGACCGGCTCGCT 6886
Db |||||
QY 1741 ATTTCCACTTACACCACTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1800
Db |||||
QY 6887 ATTTCCACTTACACCACTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 6946
Db |||||
QY 1801 TTAGCCCCCACTCTGCGTAGCATTTGCTTGGAGATACCTTTGAGTACCTGCTGACCTGCCCGGCC 1860
Db |||||
QY 6947 TTAGCCCCCACTCTGCGTAGCATTTGCTTGGAGATACCTTTGAGTACCTGCTGACCTGCCCGGCC 7006
Db |||||
QY 1861 CATACTTTTGTGATTTCTGCCAGAGTGGCGCCCTTTGGCTTTCAGGCTGCTGCTTTC 1920
Db |||||
QY 7007 CATACTTTTGTGATTTCTGCCAGAGTGGCGCCCTTTGGCTTTCAGGCTGCTGCTTTC 7066
Db |||||
QY 1921 CAGTCTACTGCTGCTGAGCTTACGCCCTTACGATGAAGTGGGTAACCTCGGAGTTG 1980
Db |||||
QY 7067 CAGTCTACTGCTGCTGAGCTTACGCCCTTACGATGAAGTGGGTAACCTCGGAGTTG 7126
Db |||||
QY 1981 TAG 1983
Db |||||
QY 7127 TAG 7129
Db |||||

RESULT 12

HPBEHV

LOCUS

HPEHEV

Hepatitis E virus genomic RNA.

D10330

VERSION

D10330.1

KEYWORDS

GI:221704

SOURCE

Hepatitis E virus

ORGANISM

Hepatitis E virus

REFERENCE

1

AUTHORS

Tam, A.W., Smith, M.M., Guerra, M.E., Huang, C.C., Bradley, D.W.,

Fry, K.E. and Reyes, G.R.

TITLE

Hepatitis E virus (HEV): molecular cloning and sequencing of the

full-length viral genome

JOURNAL

Virology 185 (1), 120-131 (1991)

MEDLINE

92024067

PUBMED

1926770

REFERENCE

2 (bases 1 to 7194)

AUTHORS

Aye, T.T., Uchida, T., Ma, X., Iida, F., Shikata, T., Ichikawa, M.,

Rikihisa, T. and Win, K.M.

TITLE

Sequence and gene structure of the hepatitis E virus isolated from

Myanmar

JOURNAL

Virus Genes 7 (1), 95-109 (1993)

MEDLINE

93227573

PUBMED

8470371

REFERENCE

3 (bases 1 to 7194)

AUTHORS

Uchida, T.

TITLE

Direct Submission

JOURNAL

Submitted (23-JAN-1992) Toshikazu Uchida, Nihon University School

of Medicine, 1st Department of Pathology; 30-1 Oyaguchi-Kaminachi,

Itabashi-Ku, Tokyo 173, Japan (Tel:81-3-3972-8111,
Fax:81-3-3972-8830)

Submitted (23-Jan-1992) to DDBJ by:

Toshikazu Uchida

Ist Department of Pathology

Nihon University School of Medicine

30-1 Oyaguchi Kamimachi

Itabashi-Ku, Tokyo 173

Japan

Phone: 81-3-3972-8111

Fax: 81-3-3972-8830.

Location/Qualifiers

1. .7194

/organism="Hepatitis B virus"

/mol_type="genomic RNA"

/isolate="HEVNE8L"

/db_xref="taxon:12461"

28. .5109

/codon_start=1

/product="ORF1"

/protein_id="BAA01172.1"

/db_xref="GI:221705"

/translation="MEAHQFIKAHGITTAEQALAAANSALANAVVRPFUSHOOIE
ILINMQPQLVREVFVWNHPIQRIHNELELYCRASGRCLIEGAPHSINDPNV
VHRCFLPVGRDQVWYAPTPTGPAANCRRSALRGLPAVDRTYCLDGFSGCNFPAETG
IALYSLHDSMSPDVAEMAFHGMHRLYAAHLHPPEVLLPPGYRTYASVLLIHDGRVY
VTEGDTAGYNHNSLRWIRTTKVTDHPLVIERVRAIGCHFVLLITAAPEPSPM
PYVPRSTEVYVRSIFPGGTPSLFPTSCSTKSTFAVPAHIMDRMLFGNLTDDQA
FCCSRMLYLGISVYVGTILVANEGWNASDALAVITAAIYLTICHOYILRTQAS
KGMRLREHQAFTITPYLSWLFKSGRDYIEGRQLEFYAQCRRWLSAGFHDLPRLV
FDESAPCRCTAIRKALSKFCFMKWLGOECTFLQPAEGVVGDDQHNEAYEGSDVD
PAESAIISISGVVPGTALOPLYALDLPAEIVARAGRLTATVKSVDGIDCELT
LGNKTRFSPVDGVALEANGPERYNSLFDASOSTWAAQPSLTYAASAGLEVRVAA
GLDHRAVPAGVSPSAGEVTAFCSALYRFNREARQSLSLTGNLWFPEGLLGLFAPP
SPGHVESANPSTLTIRTWSEVDVSSPDPDLGLMSEPSIPSRAAETPLAVLL
PPAPDPDPSPAPALDEPASGATAGAPAITHTQARHRLLTFTPDGSKVAFGLSES
TCTVLVNASINDVRPGGLCHAFYQRYPASFDASFWMRDGAAYTLTPRTIHAVAP
DYRLHNPKEALEAVRETCRSLGTAAYSLLGTGIYQVPIGPSFDAWERNRHPGDELYL
PELAARMEANPRTPTLTITEDVARTANLAJELDSANTVGBACARVTPGVQYQF
TAGVPGSKRSITQADVUDVVVPTRELNRNRGFAAFNPHTRVARTVQGRVIDE
APSLPHLLHMQAQTTHLLGDNQIPADIDFEHAGLVPAIRPDLGPTSMWHVTHRC
PADVCELLRGAYPQTTSRVLRSIFWGPAPVQKLVFQAAKAAANPGSVTVHEAQA
TYETITITADARGILQSSRAHAIVALTHTKEFVIIDAPGLLREVIGSIDAIVNFF
LAGGIEGHORPSVIRPNDNDVTLAAFPSCQISAFHOLAEELGHRVPVAAVLPP
CPLEBQGLLYLPGHATCDVSUTFEITDILVHCMAAPNQRKAVLSTLVGRYGRKLY
NASHSDRDSLARTIPAGPVQVYTICELLYELVEAMVEKQDQSAVLELDLONRDSRI
TFQXQCNKFTGETIAGKVGQGISAMSKTFCALFGPWFRAIERAKAILALLPQGVYF
DAFDDTVFSAVAAAKASVNFENDFSEDTQNNFSLGLECAIMEECGPQWLIIRLYH
LIRSAWIIQAPKESLRGFWKHSBPGFTLLMNTVMNAVITHCYDFPDQVAAFQKDD
SIVLCSYRQSPGAALVAGCLLKVDFRPIGLVAGVVVAPGLGALPDVVPFAGRLT
EKNMGGPBARASOLRLAVSDFLRLKLTNVAQMCVDVVSVRYGVPGLVHNLIGMLQAVA
DGNKAFTESVKVLDTLSILCRVE"

5106. .5477

/codon_start=1

/product="ORF3"

/protein_id="BAA01173.1"

/db_xref="GI:221706"

/translation="MNNMSFAAPMGSRPCDLGLFCCSSCFCLCPRHRPVRRLAAV
GGAAVPAVVGSTVGLIUSPSQSPIFIQPTSPSPMSPURPGLDLVFANQPDHSAPLGV
TRPSAPPLPVHVDLPLQGPGR"

5147. .7129

/codon_start=1

/product="ORF2"

/protein_id="BAA01174.1"

/db_xref="GI:221707"

/translation="MRPRPILILLMLPMLPAPPQPQSRRRRRSGSGGFGMD
RVDSOPFAIPYHTPPADVTAAAGAPRQYAPRPLGSAWROAQRPAVASRRRL
TTAGAPLTAVAHADPTPVPDVSAAILRRQYNLSTPLTSSVATGNLVLVYAAPL
SELLPLQDGTNTHIMATASNAQIRVVRATIRYRPLVNAVGGYAISISFWPQTTT
PVSVMNSITSDIRLLVQGIASELVIPISELRHYNQKWSVETSGVABEATSLGV
MLCIHGPSVNSYNTPTGALGLDPALEFLNLTGNTNTRVSRYSSTARHLRLV
ADGTALTTAATRFMKDLYFTSTNGVIGIRGIALTLFNLADTLIGLPTLISAG
GQLFYSRPVSAHGEPTVKLYTSVENAQDQKIAIPHDIDLGESRVVIQDYDNIHQEQD

RPTSPAPSPRPFSLVRANDVLWLSLTAABYDQSTYSSSTAPVYVSDSVTLNVATGAQ
AVARSLDWTKVTLDGRPLSTTQQYPKTFVLPLRGLKSFWEAGTTKAGYPYNYNTAS
DQLLVENAAAGHRVAISTVYTSILGAGPVSISAVAVLAPHSALALLBTDLTLDYPACATPD
DFCEPCRLPLGQCAQFQSTVAELQRLKMKVGTREL"

7194

polyA_site

ORIGIN

Query Match 97.6%; Score 1935; DB 14; Length 7194;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 1953; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGGCCCTATTTTGTGTCTCTCATGTCTTTTTCCTATGTCTGCCGCGCA 60

DB 5147 ATGCGACCTCGGCCCTATTTTGTGTCTCTCATGTCTTTTTCCTATGTCTGCCGCGCA 5206

QY 61 CGGCCCGGTTCAGCCGTCTGGCGCGCGGTGTGGCGCGCGAGCGGGTTCGGCGCGTGT 120

DB 5207 CGGCCCGGTTCAGCCGTCTGGCGCGCGGTGTGGCGCGCGAGCGGGTTCGGCGCGTGT 5266

QY 121 TTCTCGGGGTGACCGGGTGTGATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180

DB 5267 TTCTCGGGGTGACCGGGTGTGATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 5326

QY 181 CCCTTTCGCCCGCCGATGTTCACCGCTCGCGCGGGGTGAGACTCTGTCTTCGCAACCCCGC 240

DB 5327 CCCTTTCGCCCGCCGATGTTCACCGCTCGCGCGGGGTGAGACTCTGTCTTCGCAACCCCGC 5386

QY 241 CGAACCACTCGGCTCGGCTTGGCGTACCAAGCGCCAGCGCCCGCGCGTTCGCTCAGTCGT 300

DB 5387 CGAACCACTCGGCTCGGCTTGGCGTACCAAGCGCCAGCGCCCGCGCGTTCGCTCAGTCGT 5446

QY 301 AGACCTTACCAAGCTTGGGGCGCGCGCTTAAACCGCGGTTCGCTCCGCGCCCATGACACCCCG 360

DB 5447 AGACCTTACCAAGCTTGGGGCGCGCGCTTAAACCGCGGTTCGCTCCGCGCCCATGACACCCCG 5506

QY 361 CCAGTGCCTGTATGTTCGACTCCCGCGCGCCATCTTTGGCGCGGAGTATTAACCTATCAACA 420

DB 5507 CCAGTGCCTGTATGTTCGACTCCCGCGCGCCATCTTTGGCGCGGAGTATTAACCTATCAACA 5566

QY 421 TCTCCGCTTACCTCTTCGCTGGCGCACCGGCACCTAACCTGGTCTTTTATCGCGCCCTCTT 480

DB 5567 TCTCCGCTTACCTCTTCGCTGGCGCACCGGCACCTAACCTGGTCTTTTATCGCGCCCTCTT 5626

QY 481 AGTTCGCTTTTACCCCTTTCAGGACGCGACCAATACCCATATAATGGCCACGGAAGTCTT 540

DB 5627 AGTTCGCTTTTACCCCTTTCAGGACGCGACCAATACCCATATAATGGCCACGGAAGTCTT 5686

QY 541 AATTATGCCAGTACCGGGTTGCCGTGCCCAAAATCCGTTACCGCGCGTGTGCCCAAT 600

DB 5687 AATTATGCCAGTACCGGGTTGCCGTGCCCAAAATCCGTTACCGCGCGTGTGCCCAAT 5746

QY 601 GCTGTGGCGGTTACGCCATCTCCATCTCATCTTGGCCACAGACCAACCAACCCCGGAG 660

DB 5747 GCTGTGGCGGTTACGCCATCTCCATCTCATCTTGGCCACAGACCAACCAACCCCGGAG 5806

QY 661 TCCGTTGATATGAATTCATTAACCTCGACGAGTGTTCGTATTTTGTAGTCAGCGCCGCGCAT 720

DB 5807 TCCGTTGATATGAATTCATTAACCTCGACGAGTGTTCGTATTTTGTAGTCAGCGCCGCGCAT 5866

QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTTACACTATCGTAAACCAAGGCTGGCGCTCC 780

DB 5867 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTTACACTATCGTAAACCAAGGCTGGCGCTCC 5926

QY 781 GTCGAGACCTCTGGGGTGTGCTGAGGAGGCTACTCTGGTCTTGTATGCTTTCGCAT 840

DB 5927 GTCGAGACCTCTGGGGTGTGCTGAGGAGGCTACTCTGGTCTTGTATGCTTTCGCAT 5986

QY 841 CATGGCTCACTCGTAAATTCCTATTAACACCTATACCGGTGCCCTCGGGCTGTG 900

DB 5987 CATGGCTCACTCGTAAATTCCTATTAACACCTATACCGGTGCCCTCGGGCTGTG 6046

QY 901 GACTTTGCCCTTGAAGTTTTCGCAACCTTACCCCGGTTAACCAACCAATACCGGGGTC 960

```
Db 6047 GACTTTGCCCTTGAGCTTGAGTTTGGCAACCTTACCCCGGTAAACCAATAACGCGGTC 6106
Qy 961 TCCGGTTATTCAGCACTGCTCGCACCGGCTTGTGTCGGTGGAGCGGACTGCCAG 1020
Db 6107 TCCGGTTATTCAGCACTGCTCGCACCGGCTTGTGTCGGTGGAGCGGACTGCCAG 6166
Qy 1021 CTCACCAACCAAGGCTGTACCGCTTATGAAGGACCTCTATTATTACTAGTACTAATGGT 1080
Db 6167 CTCACCAACCAAGGCTGTACCGCTTATGAAGGACCTCTATTATTACTAGTACTAATGGT 6226
Qy 1081 GTCCGTGAGATGCGCGCGGATAGCCCTCACCCCTGTTCACCTGCTGACACTCTGCTT 1140
Db 6227 GTCCGTGAGATGCGCGCGGATAGCCCTCACCCCTGTTCACCTGCTGACACTCTGCTT 6286
Qy 1141 GCGGCTGCGGACAGAAATGATTTCTGTCGGCTGGTGGCCAGCTGTTCTACTCCCGTCC 1200
Db 6287 GCGGCTGCGGACAGAAATGATTTCTGTCGGCTGGTGGCCAGCTGTTCTACTCCCGTCC 6346
Qy 1201 GTTGTCTCAGCCAATGCGGACCGACTGTTAAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 6347 GTTGTCTCAGCCAATGCGGACCGACTGTTAAAGTTGTATACATCTGTAGAGAATGCTCAG 6406
Qy 1261 CAGATAAGGGTATTGCAATCCCGCATGACATTCAGCTCGAGAAATCTGTGTGTTATT 1320
Db 6407 CAGATAAGGGTATTGCAATCCCGCATGACATTCAGCTCGAGAAATCTGTGTGTTATT 6466
Qy 1321 CAGATTTATGATTAACCAACATGACAGATCGCGGACGCTTCTCCAGCCCATCGCG 1380
Db 6467 CAGATTTATGACCAACCAACATGACAGATCGCGGACGCTTCTCCAGCCCATCGCG 6526
Qy 1381 CCTTCTCTCTCTTCAGCTTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
Db 6527 CCTTCTCTCTCTTCAGCTTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 6586
Qy 1441 GACAGTCCACTTATGCTCTTTCAGCTGCGGCGAGTTATGTTCTGACTCTGTGACCTTG 1500
Db 6587 GACAGTCCACTTATGCTCTTTCAGCTGCGGCGAGTTATGTTCTGACTCTGTGACCTTG 6646
Qy 1501 GTTAAATGTCGACCGCGCGGAGCGTTCGCGGTCGCTCCGATTCGACCAAGGTACA 1560
Db 6647 GTTAAATGTCGACCGCGCGGAGCGTTCGCGGTCGCTCCGATTCGACCAAGGTACA 6706
Qy 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCTGCG 1620
Db 6707 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCTGCG 6766
Qy 1621 CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGACAACTAAAGCGGCTACCTTATAAT 1680
Db 6767 CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGACAACTAAAGCGGCTACCTTATAAT 6826
Qy 1681 TATAACACCACTGCTAGCGACCAACTGCTTTGTGAGAAATCCGCGCGGACCGGTCGCT 1740
Db 6827 TATAACACCACTGCTAGCGACCAACTGCTTTGTGAGAAATCCGCGCGGACCGGTCGCT 6886
Qy 1741 ATTTTCCACTTACACCACTAGCTGGGTGCTGGTCCGCTCTCCATTTCTGCGGTTGCGGTT 1800
Db 6887 ATTTTCCACTTACACCACTAGCTGGGTGCTGGTCCGCTCTCCATTTCTGCGGTTGCTGTT 6946
Qy 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGATACCTTGAGACTACCTGCGCGGCG 1860
Db 6947 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGATACCTTGAGACTACCTGCGCGGCG 7006
Qy 1861 CATACTTTTGTATGATTTCTGCCAGAGTGCAGCCCTCTGCGCTTACAGGCTCGCTTTC 1920
Db 7007 CATACTTTTGTATGATTTCTGCCAGAGTGCAGCCCTCTGCGCTTACAGGCTCGCTTTC 7066
Qy 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGTGAAGGTGGGTGAGTAACTCGGGAGTTG 1980
Db 7067 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGTGAAGGTGGGTGAGTAACTCGGGAGTTA 7126
Qy 1981 TAG 1983
Db 7127 TAG 7129
```

RESULT 13

```
AF124407
LOCUS      1983 bp      RNA      linear      VRL 24-JUL-1999
DEFINITION Hepatitis E virus isolate AKI-90 ORF-2 gene, complete cds.
ACCESSION  AF124407
VERSION     AF124407.1  GI:5579229
KEYWORDS
SOURCE
  ORGANISM  Hepatitis E virus
            Hepatitis E virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Hepatitis
            E-like viruses.
REFERENCE   1 (bases 1 to 1983)
AUTHORS    Arankalle,V.A., Paranjape,S., Emerson,S.U., Purcell,R.H. and
            Walimbe,A.M.
TITLE      Phylogenetic analysis of hepatitis E virus isolates from India
            (1976-1993)
JOURNAL    J. Gen. Virol. 80 (Pt 7), 1691-1700 (1999)
MEDLINE    9350000
PUBMED     10423137
REFERENCE   2 (bases 1 to 1983)
AUTHORS    Arankalle,V.A., Paranjape,S., Emerson,S.U., Purcell,R.H. and
            Walimbe,A.M.
TITLE      Direct Submission
JOURNAL    Submitted (28-JAN-1999) Hepatitis, National Institute of Virology,
            20-A Dr. Ambedkar Road, Pune, Maharashtra 411001, India
            Location/Qualifiers
FEATURES
  source
    1..1983
       /organism="Hepatitis E virus"
       /mol_type="genomic RNA"
       /isolate="AKI-90"
       /db_xref="taxon:12461"
       /country="India"
    1..1983
       /codon_start=1
       /product="ORF-2"
       /protein_id="AAD45493.1"
       /db_xref="GI:5579230"
       /translation="MRPRPILLLLMFLPMLPAPPPGQPSRRRRSGSGGGPWGD
            RVDSPALPIYHPTNFPADVTAAAGARRINOPARPLGSAWRDQAPPAASRRRP
            TTAGAAPLAVAPADHTPPVDVDSRGAIRLNQYNLSTPLTSVATGNTLVYAAPL
            SPPLPQDGTNTTHIMATSNYAQRVARATIRPLVPNAWGSVETSGVAEEATGTV
            PSLVDMSITSDVRLVQPGIASELVIPSELRHYNQWNRVETSGVAEEATGTV
            MLCIHGSPNSVNTPTGALGILDEPLENTPGNNTVSVSVSTARRLSRG
            AGTAEITTAATREMKDLFFSTNGVGEIGRGIATLNLATLLGLPTELLISAG
            GQLFYSRPVWSANGEPTVKLYTSVENAQDDKGIAPHDIDLGESRVVIQYDNOHEQ
            RPTSPAPGRPFVLRANDVLMLSLTAAYDQSTGSGTSPVTVSDSVTLNVATGAQ
            AVARSLDWTKTLDGRPLSTIQOYSKTFVFLPLRKLFSWEAGTTKAGYPYNTTAS
            DOLLVENAAGHRVAISTYTTSLGAGPVSISAVAVAFAPHSALALLEDTLDYPARAHTFD
            DFCPECRPLGQCAFQSVAEIQLRLKMKVGKTRFL"
```

ORIGIN

```
Query Match      96.5%; Score 1914.2; DB 14; Length 1983;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1940; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 ATGCGCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCGCTATGTCGCCGCGCA 60
Db 1 ATGCGCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCGCTATGTCGCCGCGCA 60
Qy 61 CGCGCCGCTCAGCGCTGTCGCGCGCTGCTGCGCGCGAGCGCGGTCCTCGCGGCTG 120
Db 61 CGCGCCGCTCAGCGCTGTCGCGCGCTGCTGCGCGCGAGCGCGGTCCTCGCGGCTG 120
Qy 121 TTTCTGGGGTGACCGGTTGATTCAGCCCTTCGCAATCCCTTATATTCATCAACCAAC 180
Db 121 TTTCTGGGGTGACCGGTTGATTCAGCCCTTCGCAATCCCTTATATTCATCAACCAAC 180
Qy 181 CCCTTCGCCCCCGATCTCAGCGCTGCGCGGCGGCTGACCTGCTGTCGACACCGCC 240
Db 181 CCCTTCGCCCCCGATCTCAGCGCTGCGCGGCGGCTGACCTGCTGTCGACACCGCC 240
```

241 CGAACCTCGGCTCCGCTTGGCTGACACAGGCCCGAGCGCCCGCGCTTGCCTCACTCGT 300
|||||
Db |||||
241 CGAACCTCGGCTCCGCTTGGCTGACACAGGCCCGAGCGCCCGCGCTTGCCTCACTCGT 300
|||||
Qy |||||
301 AGACCTACACAGCTGGGGCGCGCGCTTAACCGGGTGGCTCGCGGCCATGACACCCCG 360
|||||
Db |||||
301 AGACCTACACAGCTGGGGCGCGCGCTTAACCGGGTGGCTCGCGGCCATGACACCCCG 360
|||||
Qy |||||
361 CCAGTGCTGATGTCGACTCCGCGGGCGCACTTTGGCGCGGAGTATTAACCTATCAACA 420
|||||
Db |||||
361 CCAGTGCTGATGTCGACTCCGCGGGCGCACTTTGGCGCGGAGTATTAACCTATCAACA 420
|||||
Qy |||||
421 TCTCCCTTACCTCTTCCTGGTGGCCACCGGCACTAACCTGGTCTTTATGCGCCCTCTT 480
|||||
Db |||||
421 TCTCCCTTACCTCTTCCTGGTGGCCACCGGCACTAACCTGGTCTTTATGCGCCCTCTT 480
|||||
Qy |||||
481 AGTCGGCTTTACCTCTTACGAGCGGCACTAATACCTATATATGCGGCAAGGTTCT 540
|||||
Db |||||
481 AGTCGGCTTTACCTCTTACGAGCGGCACTAATACCTATATATGCGGCAAGGTTCT 540
|||||
Qy |||||
541 AATTATGCCAGTACGGGTTGCCCGTGCACCAATCCGTTACCGCGCGTGCCTCCAT 600
|||||
Db |||||
541 AATTATGCCAGTACGGGTTGCCCGTGCACCAATCCGTTACCGCGCGTGCCTCCAT 600
|||||
Qy |||||
601 GCTGTCGGCGTTACGCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCCGGAG 660
|||||
Db |||||
601 GCTGTCGGCGTTACGCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCCGGAG 660
|||||
Qy |||||
661 TCCGTTGATATGAATTCAATAACCTCGACGGATGTTGATTTATGTCAGCCCGGCATA 720
|||||
Db |||||
661 TCCGTTGATATGAATTCAATAACCTCGACGGATGTTGATTTATGTCAGCCCGGCATA 720
|||||
Qy |||||
721 GCCTCGAGCTGTGATCCCAAGTCAGCGCTACACTATCGTAACCAAGGTCGGCTCC 780
|||||
Db |||||
721 GCCTCGAGCTGTGATCCCAAGTCAGCGCTACACTATCGTAACCAAGGTCGGCTCC 780
|||||
Qy |||||
781 GTCGAGACTCTCGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
|||||
Db |||||
781 GTCGAGACTCTCGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
|||||
Qy |||||
841 CATGGCTACTCGTAAATTCCTATATAACCTATACCGGTATACCGGTGCGCTGTGG 900
|||||
Db |||||
841 CATGGCTACCGGTAAATTCCTATATAACCTATACCGGTATACCGGTGCGCTGTGG 900
|||||
Qy |||||
901 GACTTTGCCCTGAGCTGAGTTTCGACACTTACCCCGGTAAACCAATACCGGGTCC 960
|||||
Db |||||
901 GACTTTGCCCTGAGCTGAGTTTCGACACTTACCCCGGTAAACCAATACCGGGTCC 960
|||||
Qy |||||
961 TCCCGTTATTCAGCACTCTCGCCACCGCTTCTGTCGGTGGCGGAGCGGACTGCGGAG 1020
|||||
Db |||||
961 TCCCGTTATTCAGCACTCTCGCCACCGCTTCTGTCGGTGGCGGAGCGGACTGCGGAG 1020
|||||
Qy |||||
1021 CTCACACACAGCTGCTACCGCTTTATGAAGGAGCTCTATTTACTAGTACTAATGGT 1080
|||||
Db |||||
1021 CTCACACACAGCTGCTACCGCTTTATGAAGGAGCTCTATTTACTAGTACTAATGGT 1080
|||||
Qy |||||
1081 GTCGGTGGATCGCGCGGGATAGCCCTACCCCTGTTCACCTGCTGACACTCTGCTT 1140
|||||
Db |||||
1081 GTCGGTGGATCGCGCGGGATAGCCCTACCCCTGTTCACCTGCTGACACTCTGCTT 1140
|||||
Qy |||||
1141 GCGCGCTTCCGACAGAAATTTGTTCTGTCGGTGGTGGCGAGCTTCTACTCCCGTCCC 1200
|||||
Db |||||
1141 GCGCGCTTCCGACAGAAATTTGTTCTGTCGGTGGTGGCGAGCTTCTACTCCCGTCCC 1200
|||||
Qy |||||
1201 GTTGTCTACGCAATGGCGAGCGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
|||||
Db |||||
1201 GTTGTCTACGCAATGGCGAGCGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
|||||
Qy |||||
1261 CAGGATAAGGTTATTCATATCCGATGACATGACCTCGGAAATCTCGTGTGTTATT 1320
|||||
Db |||||
1261 CAGGATAAGGTTATTCATATCCGATGACATGACCTCGGAAATCTCGTGTGTTATT 1320
|||||
Qy |||||
1321 CAGGATTATGATAACCAATGAACAGATCGCGCGAGCGCTTCTCCAGCCCCATCGCGC 1380
|||||

Db |||||
1321 CAGGATTATGATAACCAACATGAACAGACCGCGCGCTTCTCCAGCCCCATCGCGC 1380
|||||
Qy |||||
1381 CTTTCTCTGCTTCGAGCTAAATGATGCTTTGGCTCTCTCTCACCCTGCGAGTAT 1440
|||||
Db |||||
1381 CTTTCTCTGCTTCGAGCTAAATGATGCTTTGGCTCTCTCTCACCCTGCGAGTAT 1440
|||||
Qy |||||
1441 GACCACTCACCTATGCTTCGACTGCGCGCTTATGTTTCTGACTCTGCTGACTTG 1500
|||||
Db |||||
1441 GACCACTCACCTATGCTTCGACTGCGCGCTTATGTTTCTGACTCTGCTGACTTG 1500
|||||
Qy |||||
1501 GTTAATGTTGCGACCGCGCGCGCTTGCCTGCTGCTGATTTGACCAAGGTACA 1560
|||||
Db |||||
1501 GTTAATGTTGCGACCGCGCGCGCTTGCCTGCTGCTGATTTGACCAAGGTACA 1560
|||||
Qy |||||
1561 CTTGAGGTCGCGCCCTCTCCACCATCAGAGTACTCGAAGACCTTCTTTGCTCGCG 1620
|||||
Db |||||
1561 CTTGAGGTCGCGCCCTCTCCACCATCAGAGTACTCGAAGACCTTCTTTGCTCGCG 1620
|||||
Qy |||||
1621 CTTCCGCGTAACTCTCTTTCTGGAGGAGGACCAACTAAAGCCGGGTACCTTATAAT 1680
|||||
Db |||||
1621 CTTCCGCGTAACTCTCTTTCTGGAGGAGGACCAACTAAAGCCGGGTACCTTATAAT 1680
|||||
Qy |||||
1681 TATAACACACTGCTAGCGACCACTGCTTGTGAGAAATGCCCGCGGACACCGGTGCT 1740
|||||
Db |||||
1681 TATAACACACTGCTAGCGACCACTGCTTGTGAGAAATGCCCGCGGACACCGGTGCT 1740
|||||
Qy |||||
1741 ATTTCACCTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
|||||
Db |||||
1741 ATTTCACCTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
|||||
Qy |||||
1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCGCGCGC 1860
|||||
Db |||||
1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCGCGCGC 1860
|||||
Qy |||||
1861 CATACTTTTGTATGATTTCTGCGCCAGAGTCCGCCCTTTGCGCTTCAAGGGCTGCGCTTTC 1920
|||||
Db |||||
1861 CATACTTTTGTATGATTTCTGCGCCAGAGTCCGCCCTTTGCGCTTCAAGGGCTGCGCTTTC 1920
|||||
Qy |||||
1921 CAGTCTACTGCTGCTGAGCTTACAGCGCTTAAAGTGAAGTGGGTAAAGTTCGGAGTTG 1980
|||||
Db |||||
1921 CAGTCTACTGCTGCTGAGCTTACAGCGCTTAAAGTGAAGTGGGTAAAGTTCGGAGTTG 1980
|||||
Qy |||||
1981 TAG 1983
|||
Db |||||
1981 TAG 1983
|||

RESULT 14
AF051830

LOCUS AF051830 7199 bp RNA linear VRL 18-DEC-1998
DEFINITION Hepatitis E virus strain TK15/92, complete genome.

ACCESSION AF051830

VERSION AF051830.1 GI:4033730

KEYWORDS

SOURCE Hepatitis E virus

ORGANISM Hepatitis E virus

Viruses; ssRNA positive-strand viruses, no DNA stage; Hepatitis E-like viruses.

REFERENCE 1 (bases 1 to 7199)

AUTHORS Gouvea V., Snellings N., Popek M.J., Longer C.F. and Innis B.L.

TITLE Hepatitis E virus: complete genome sequence and phylogenetic analysis of a Nepali isolate

JOURNAL Virus Res. 57 (1), 21-26 (1998)

MEDLINE 99049628

PUBMED 9833882

REFERENCE 2 (bases 1 to 7199)

AUTHORS Gouvea V.

TITLE Direct Submission

JOURNAL Submitted (02-MAR-1998) Department of Virus Diseases, Walter Reed

Army Institute of Research, 14th and Dahlia St., Bldg. 40,

Washington DC 20307-5100, USA

FEATURES Location/Qualifiers

source	1. .7199	organism="Hepatitis E virus"	Query Match	95.6%;	Score 1896.6;	DB 14;	Length 7199;
	/mol_type="genomic RNA"		Best Local Similarity	97.3%;	Pred. No. 0;		
	/strain="TK15/92"		Matches 1929;	Conservative	0;	Mismatches	54;
	/db_xref="taxon:12461"					Indels	0;
	/country="Nepal"					Gaps	0;
	/note="isolated from the bile of experimentally infected cynomolgous monkey"						
CDS	28. .5109		Qy	1	ATCGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA	60	
			Db	5147	ATCGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA	5206	
			Qy	61	CCGCGCGCTCAGCCGCTGCGCCCGCTGCTGGCGGGGAGCGCGGTTCCGGGGTGGT	120	
			Db	5207	CCGCGCGCTCAGCCGCTGCGCCCGCTGCTGGCGGGGAGCGCGGTTCCGGGGTGGT	5266	
			Qy	121	TTCTGGGGTGACCGGGTGTGATTCTCAGCCCTTCGCAATCCCTTATATTCATCCAAACCAAC	180	
			Db	5267	TTCTGGGGTGACCGGGTGTGATTCTCAGCCCTTCGCAATCCCTTATATTCATCCAAACCAAC	5326	
			Qy	181	CCCTTCGCGCCCGATGTCAACCGCTGCGCGCGGGCTGGACCTCGTGTGCGCCAAACCGGC	240	
			Db	5327	CCCTTCGCGCCCGATGTCAACCGCTGCGCGCGGGCTGGACCTCGTGTGCGCCAAACCGGC	5386	
			Qy	241	CGACCACTCGGCTCCGCTTGGCGTGGACCGCCGAGCGCGCGCGCTTGCTTGCCTCACGCTGT	300	
			Db	5387	CGACCACTCGGCTCCGCTTGGCGTGGACCGCCGAGCGCGCGCGCTTGCTTGCCTCACGCTGT	5446	
			Qy	301	AGACCTTACCACAGCTGGGCGCGCGCTAACCGCGCTCGCTCGGCGCCCATGACACCCCG	360	
			Db	5447	AGACCTTACCACAGCTGGGCGCGCGCTAACCGCGCTCGCTCGGCGCCCATGACACCCCG	5506	
			Qy	361	CCAGTGCCTGATGTGCACTCCCGCGCGCCATCTTGGCGCGGAGTAACTATCAACA	420	
			Db	5507	CCGCTGCCTGATGTGCGCTCCCGCGCGCCATCTTGTGCGCGGAGTAACTATCAACA	5566	
			Qy	421	TCTCCCTTACCTCTTCCGTTGGCGACCGGACCTAACTGCTGTTTATGCGCGCTCTT	480	
			Db	5567	TCTCCCTTACCTCTTCCGTTGGCGACCGGACCTAACTGCTGTTTATGCGCGCTCTT	5626	
			Qy	481	AGTCCGCTTTACCTCTTCCGTTGGCGACCGGACCTAACTGCTGTTTATGCGCGCTCTT	540	
			Db	5627	AGTCCGCTTTACCTCTTCCGTTGGCGACCGGACCTAACTGCTGTTTATGCGCGCTCTT	5686	
			Qy	541	AATATGCGCGAGTACCGGGTGGCGGCGCAAACTCGGTTACCGCGCGCTGTTGCCCAAT	600	
			Db	5687	AATATGCGCGAGTACCGGGTGGCGGCGCAAACTCGGTTACCGCGCGCTGTTGCCCAAT	5746	
			Qy	601	GCTGTGCGGGTTACCGCATCTCCATCTCATCTGCTGCGGACAGACCAACCCCGACG	660	
			Db	5747	GCTGTGCGGGTTACCGCATCTCCATCTCATCTGCTGCGGACAGACCAACCCCGACG	5806	
			Qy	661	TCCGTTGATATGAATTAATCAATAACCTCGACGGATGTTTCTGATTTTATGTCAGCGCGGATA	720	
			Db	5807	TCCGTTGATATGAATTAATCAATAACCTCGACGGATGTTTCTGATTTAGTCCAGCGCGGATA	5866	
			Qy	721	GCTCTGAGCTTGTGATATCCCAAGTACGCGCTACACTTCGTAAACAGGCTGGCGCTCC	780	
			Db	5867	GCTCTGAGCTTGTGATATCCCAAGTACGCGCTACACTTCGTAAACAGGCTGGCGCTCC	5926	
			Qy	781	GTGAGACCTCTCGGGTGGCTGAGGAGGAGGCTACCTCTGCTGTTTATGCTTTGATA	840	
			Db	5927	GTGAGACCTCTCGGGTGGCTGAGGAGGAGGCTACCTCTGCTGTTTATGCTTTGATA	5986	
			Qy	841	CATGGCTCACTCGTAAATTCCTATCTATCTAATACACCTTATACCGGTCCTCGGGCTGTTG	900	
			Db	5987	CATGGCTCACTCGTAAATTCCTATCTAATACACCTTATACCGGTCCTCGGGCTGCTG	6046	
			Qy	901	GACTTGGCCCTTGAGCTTGTGATTTCCGACCTTACCCCGGTAACCACTATACCGGGTC	960	
			Db	6047	GACTTGGCCCTTGAGCTTGTGATTTTCGACCTTACCCCGGTAACCACTATACCGGGTC	6106	
			Qy	961	TCCCGTTATTCACGACTGTCTGCCACCGCTTCTGCTGGGTCGCGAGCTGCGGAG	1020	
			Db	6107	TCCCGTTATTCACGACTGTCTGCCACCGCTTCTGCTGGGTCGCGAGCTGCGGAG	6166	

GGAAYPAVVGVTGLILSPSQSIPFIQPTTPSPMPLRPLGLDLVFPANPDHSAPLGV
TRPSAPLPHVVDLPLQGPGR"
5036..7078
/note="ORF2"
/codon_start=1
/product="structural protein"
/protein_id="AAG16766.1"
/db_xref="GI:10336835"
translation="MRPRPILLLMLPMLPAPPQPSRRRRSGSGSGGFWD
RVDSQPAIFIPTFFADVDTAAGAGPRVQPARPLGSAMRDQAPAAARRRP
ATGAPLTAFAHPTDFPDVDSRGAIRLNLSLTSVATGNLVLAAAPL
SPLLQDGNTHIMATASNAQYVARATIRYPLVFNAGVGIASISFWPQTIT
PTSDMNSITSDIRILVQPGIASLWIPSERLHYRNOQWRSVETSGABEEATSLV
MLCIGSPVNSYNTPTVGTALGLDPALEFRNLTPGNTNTRVSYSSARHLRRLG
ADCTAELTTTAAFRMKDLVFTSTNGVEIGRGIALTLNLADTLGLLPTLILSAG
GQIFYSRPVVSAENGEPVKLYTSVENAQDKGIALPHDIDLGRSVVQIDYDNOHQD
RPTSPAPSPFVSLANDVLMLSLTAAYEQDYSSTGTPVYVSDSTLVNLTVAAG
AVARSLDWTKTLIDGRPLSTIQYKSTFVPLRGLKSFWEAGTTKAGYVNYNTAS
DQLLVENAGHRVAISYITTSLSGAGPVVSI SAVAVLAPHSA LALLEDTLDYPARAHTPD
DFCPECRPLGLQCAFOSTVAELQRLMKVGTREL"

ORIGIN

Query Match 95.2%; Score 1888.6; DB 14; Length 7143;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1924; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCTATTGTTGCTGCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
Db |||||
QY 61 CGCGCCGCTCAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db |||||
QY 121 TTCTGGGGTACCGGGTGAATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db |||||
QY 5216 TTCTGGGGTGACCGGGTGAATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 5275
QY 181 CCCTTGCCTCCCGGATGTCACGCTGCGCGCGGGCTGGACCTCGTTCGCCAACCGCC 240
Db |||||
QY 5276 CCCTTGCCTCCCGGATGTCACGCTGCGCGCGGGCTGGACCTCGTTCGCCAACCGCC 5335
QY 241 CGACCACTCGGCTCGGCTGGCGTGACAGGCCCGCGCGCGCGCGCTGCTCAGCGTGT 300
Db |||||
QY 5336 CGACCACTCGGCTCGGCTGGCGTGACAGGCCCGCGCGCGCGCGCTGCTCAGCGTGT 5395
QY 301 AGACCTACCAAGTGGGGCGGCGCGCTAACCGCGTGTCTCGGCGCGATGACACCGG 360
Db |||||
QY 5396 AGACCTGCGCAGCTGGGGCGGCGCGCTAACCGCGTGTCTCGGCGCGATGACACCGG 5455
QY 361 CCAAGTGCCTGATGTCAGCTCCCGGGCGGCATCTTGGCGCGCGAGTAACTTATCAACA 420
Db |||||
QY 5456 CCAAGTGCCTGATGTCAGTTCGCGGGCGGCATCTTGGCGCGCGAGTAACTTATCAACA 5515
QY 421 TCTCCGCTTACTCTTCCGTTGGCGACCGGCACTAACTGTTCTTTATGCGCGCCCTCTT 480
Db |||||
QY 5516 TCTCCGCTTACTCTTCCGTTGGCGACCGGCACTAACTGTTCTTTATGCTGCTCCCTT 5575
QY 481 AGTCCGCTTACTCCCTTCAGGACGCGCACTAACCTATATATGCGCGAGCTTCT 540
Db |||||
QY 5576 AGCCCACTTCTACCCCTTCAGGACGCGCACTAACTATATATGCGCGAGCTTCT 5635
QY 541 AATTATGCCAGTACCGGGTTCGCCGTGCGCACAACTTACCGCGCGCTGGTCCCAAT 600
Db |||||
QY 5636 AATTATGCCAGTACCGGGTTCGCCGTGCGCACAACTTACCGCGCGCTGGTCCCAAT 5695
QY 601 GCTGTGGCGGGTTACGCACTCTCCATCTATTCTGGCCACAGACCAACCAACCGGAGG 660
Db |||||
QY 5696 GCTGTGGCGGGTTACGCACTCTCCATCTCATTTTGGCCACAGACCAACCAACCGGAGG 5755
QY 661 TCCGTTGATATGAATCAATAAATCTCGAGGATGTTGATTTTCTAGTCCAGCGCGGAT 720
Db |||||
QY 5756 TCCGTTGATATGAATCAATAAATCTCGAGGATGTTGATTTTCTAGTCCAGCGCGGAT 5815
Db |||||

QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGCGGTCC 780
Db |||||
QY 5816 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGCGGTCC 5875
QY 781 GTCGAGACCTCTCGGGTGGCTGAGGAGGAGGCTACCTCTGCTGTTGTTATGCTTTGCATA 840
Db |||||
QY 5876 GTCGAGACCTCTCGGGTGGCTGAGGAGGAGGCTACCTCTGCTGTTGTTATGCTTTGCATA 5935
QY 841 CATGGCTCACTCGTAATATCTTATACATACACCTTATACCGGTGCCCTCGGGCTGTG 900
Db |||||
QY 5936 CATGGCTCACTCGGTAAATCTTACATTAATACACCTTATACCGGTGCCCTCGGGCTGTG 5995
QY 901 GACTTTGCCCTTCAGCTTGGAGTTTCCCAACTTACCCCGGTAAACCAATACCGGAGT 960
Db |||||
QY 5996 GACTTTGCCCTTCAGCTTGGAGTTTCCCAACTTACCCCGGTAAACCAATACCGGAGT 6055
QY 961 TCCCGTTATTTCAGCACTGCTCGCCACCGCTTTCGTCGCGGTGGAGCGGAGTGCAG 1020
Db |||||
QY 6056 TCCCGTTATTTCAGCACTGCTCGCCACCGCTTTCGTCGCGGTGGAGCGGAGTGCAG 6115
QY 1021 CTCACCAACAGGCTGCTACCGCTTTATGAAGGACCTCTATTCTAGTACTAATGGT 1080
Db |||||
QY 6116 CTCACCAACAGGCTGCTACCGCTTTTATGAAGGACCTCTATTCTAGTACTAATGGT 6175
QY 1081 GTCGGTGAGATCGCGCGGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
Db |||||
QY 6176 GTCGGTGAGATCGCGCGGGATAGCTCTCACCTGTTCAACCTTGCTGACACTCTGCTT 6235
QY 1141 GCGCGCTCGCGACAGAAATGATTTCGTCGGCTGGTGGCCAGCTGTTCTACTCCGTC 1200
Db |||||
QY 6236 GGTGGCTCGCGACAGAAATGATTTCGTCGGCTGGTGGCCAGCTGTTCTACTCTCGTCC 6295
QY 1201 GTTGTCTCAGCAATCGCGAGCGAGCTGTTAAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Db |||||
QY 6296 GTGCTCTCAGCAATCGCGAGCGAGCTGTTAAAGCTGTATACATCTGTAGAGAAATGCTCAG 6355
QY 1261 CAGGATAAGGATTCGAATCCCGCATGACATTCGAGAACTCGGAGAACTCGTGTGGTTATT 1320
Db |||||
QY 6356 CAGGATAAGGATTCGAATCCCGCATGACATTCGAGAACTCGGAGAACTCGTGTGGTTATT 6415
QY 1321 CAGGATAATGATAACCAACATGAACAGATCGGCGCGAGCTTCTCAGCGCCCATCGCGC 1380
Db |||||
QY 6416 CAGGATAATGATAACCAACATGAACAGAGCGCGCGAGCGCTTTCGCGAGCTCCATCGCGC 6475
QY 1381 CCTTTCTCTGTCTTCTGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT 1440
Db |||||
QY 6476 CCTTTCTCTGTCTTCTGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT 6535
QY 1441 GACAGTCCACTTATGGCTCTTCTGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db |||||
QY 6536 GACCAATCCACTTATGGCTCTTCTGACTGGCCAGTTTATGTTCTGACTCTGTGACCTTG 6595
QY 1501 GTTAAATGTTGCGACCGCGCGAGCGCTTTCGCGGTGCGCTCGAATGGACCAAGGTGACA 1560
Db |||||
QY 6596 GTTAAATGTTGCGACCGCGCGAGCGCTTTCGCGGTGCGCTCGAATGGACCAAGGTGACA 6655
QY 1561 CTTGAGCGGTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCTGCGG 1620
Db |||||
QY 6656 CTTGAGCGGTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCTGCGG 6715
QY 1621 CTTCCGCGGTGAAGCTCTCTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCCCTTATAAT 1680
Db |||||
QY 6716 CTTCCGCGGTGAAGCTCTCTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCCCTTATAAT 6775
QY 1681 TATAACCACTGCTAGCGACCAACTGCTGTGCGAAGATGCGCGCGGACCGCGGTGCTG 1740
Db |||||
QY 6776 TATAACCACTGCTAGCGACCAACTGCTGTGCGAAGATGCGCGCGGACCGCGGTGCTG 6835
QY 1741 AATTCCACCTTACCACTAGCGGTGCTGCTGCGGTCTCCATTTCTCGGTTGCGGTT 1800
Db |||||
QY 6836 AATTCCACCTTACCACTAGCGGTGCTGCTGCGGTCTCCATTTCTCGGTTGCTGCTGTT 6895
Db |||||

Qy	1801	TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGACTACCTGCCCGCGCC	1860
Db	6896	TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGACTACCTTGCGCGCC	6955
Qy	1861	CATACTTTTGATGATTTCTGCCAGAGTGCAGCGCCCTTGCGCTTCAGGGCTGCGTTTC	1920
Db	6956	CATACTTTTGATGATTTCTGCCAGAGTGCAGCGCCCTTGCGCTTCAGGGCTGCGTTTC	7015
Qy	1921	CAGTCTACTGCTGAGCTTCAGCGCCTTAAGATGAAGTGGTAAACTCGGAGTTG	1980
Db	7016	CAGTCTACTGCTGAGCTTCAGCGCCTTAAGATGAAGTGGTAAACTCGGAGTTA	7075
Qy	1981	TAG 1983	
Db	7076	TAG 7078	

Search completed: August 21, 2004, 15:14:59
Job time : 7653.28 secs

Blank sheet

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 08:40:47 ; Search time 765.579 Seconds
(without alignments)
11003.666 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129
Perfect score: 1983
Sequence: 1 ATGCGCCCTCGGCTATTATTT.....GTAAACTCGGAGTGTAG 1983

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1983	100.0	1983	2	AAT96959	Aat96959 Hepatitis
2	1983	100.0	2049	2	AAT27107	Aat27107 Hepatitis
3	1983	100.0	2094	2	AAQ47129	Aaq47129 HEV third
4	1983	100.0	2094	2	AAQ46813	Aaq46813 Burma str
5	1983	100.0	7195	2	AAV66321	Aav66321 ET-NANB (
6	1983	100.0	7195	2	AAV54729	Aav54729 DNA seque
7	1983	100.0	7195	3	AAQ99259	Aaq99259 Hepatitis
8	1983	100.0	7195	4	AAF83495	Aaf83495 ET-NANB v
9	1983	100.0	7195	6	AAAL50386	Aal50386 Hepatitis
10	1978.2	99.8	7195	9	ADD24374	Add24374 Forward s
11	1970.2	99.4	7195	2	AAQ14412	Aaq14412 Hepatitis
12	1949.4	98.3	1983	2	AAV61688	Aav61688 Hepatitis
13	1949.4	98.3	7194	2	AAV61690	Aav61690 Hepatitis
14	1805.4	91.0	1984	2	AAQ86592	Aaq86592 Hepatitis
15	1805.4	91.0	1990	6	ABL60053	Abi60053 Hepatitis
16	1805.4	91.0	1990	9	ADB97808	Adb97808 HEV ORF2
17	1805.4	91.0	2054	4	AAAS01448	Aas01448 Hepatitis
18	1793.4	90.4	1989	6	ABL60054	Abi60054 Hepatitis
19	1793.4	90.4	1989	9	ADB97809	Adb97809 HEV relat
20	1791	90.3	7204	8	ADA50063	Ada50063 SK-HEV-2
21	1789.4	90.2	7204	8	ADA50062	Ada50062 SK-HEV-3
22	1789.4	90.2	7204	8	ADA50065	Ada50065 Hepatitis
23	1787.8	90.2	7158	2	AAT27394	Aat27394 Hepatitis

24	1787.8	90.2	7168	2	AAQ45197	Aaq45197 HEV strai
25	1787.8	90.2	7168	2	AAV71604	Aav71604 Hepatitis
26	1787.8	90.2	7204	8	ADA50064	Ada50064 Hepatitis
27	1647	83.1	1647	2	AAT27109	Aat27109 Hepatitis
28	1617.2	81.6	1647	2	AAV61686	Aav61686 Hepatitis
29	1506.4	76.0	1643	2	AAQ27436	Aaq27436 DNA encod
30	1384.6	69.8	2058	2	AAT27108	Aat27108 Hepatitis
31	1368.6	69.0	1980	2	AAT96960	Aat96960 Hepatitis
32	1368.6	69.0	2100	2	AAQ47130	Aaq47130 HEV third
33	1368.6	69.0	7171	2	AAQ14413	Aaq14413 Composite
34	1368.6	69.0	7171	2	AAV66322	Aav66322 ET-NANB (
35	1368.6	69.0	7171	2	AAV54730	Aav54730 Composite
36	1368.6	69.0	7171	3	AAQ99260	Aaq99260 Hepatitis
37	1368.6	69.0	7171	4	AAF83496	Aaf83496 Mexican s
38	1368.6	69.0	7171	6	AAAL50387	Aal50387 Hepatitis
39	1368.6	69.0	7171	9	ADD24378	Add24378 Hepatitis
40	1368.6	69.0	7180	9	ADE06697	Ade06697 Hepatitis
41	1367.6	69.0	7230	9	ADE06698	Ade06698 Hepatitis
42	1367	68.9	2100	2	AAQ46814	Aaq46814 Mexico st
43	1347.6	68.0	7228	9	ADE06696	Ade06696 Hepatitis
44	1346.8	67.9	7233	9	ADE06694	Ade06694 Hepatitis
45	1346	67.9	7234	9	ADE06660	Ade06660 Hepatitis

ALIGNMENTS

RESULT 1
AAT96959
ID AAT96959 standard; DNA; 1983 BP.
XX
AC AAT96959;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 26-FEB-1998 (first entry)
XX
DE Hepatitis E virus Burma strain ORF2.
XX
KW Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody;
KW diagnosis; HEV; ds.
XX
OS Hepatitis E virus; - Burma strain.
XX
FH Key Location/Qualifiers
FT CDS 1..1983
FT /*tag= a
FT /product= "ORF2_protein"
XX
PN US5686239-A.
XX
PD 11-NOV-1997.
XX
PF 09-MAY-1994; 94US-00240049.
XX
PR 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-0036672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 17-JAN-1991; 91US-00681078.
PR 20-APR-1992; 92US-00822335.
PR 01-MAY-1992; 92US-00870985.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Tam AW, Reyes GR, Yarbough PO;
XX
DR WPI; 1997-558132/51.
XX
P-PSDB; AAW35826.
PT Diagnosis of hepatitis E virus Burma and Mexico strain infection - by

immunoassay with hepatitis E virus peptide antigens.

Disclosure; Col 23-24; 36pp; English.

A method has been developed for detecting hepatitis E virus (HEV) antibodies (Ab). The method comprises: (a) reacting a serum sample with a HEV peptide antigen; and (b) examining the peptide for the presence of bound Ab, where the presence of bound Ab indicates the presence of HEV Ab. The present sequence encodes the protein from the open reading frame, ORF2, from HEV Burma strain. The method can be used to diagnose infection with the enterically transmitted non-A/non-B viral hepatitis agent HEV, specifically the HEV Burma and Mexico strains. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 17-OCT-2003 to standardise OS field)

Sequence 1983 BP; 340 A; 665 C; 486 G; 492 T; 0 U; 0 Other;

Query Match 100.0%; Score 1983; DB 2; Length 1983;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGCTGCGCGCCA	60
DB	1	ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGCTGCGCGCCA	60
QY	61	CGCGCGGTCAGCGCTCGCGCGCTCGTGGCGCGCAGCGCGGTTCCGCGGTGGT	120
DB	61	CGCGCGGTCAGCGCTCGCGCGCTCGTGGCGCGCAGCGCGGTTCCGCGGTGGT	120
QY	121	TTCTGCGGGTGACCGGGTGATTTCTAGCCCTTCGCAATCCCTPATATTCATCAACCAAC	180
DB	121	TTCTGCGGGTGACCGGGTGATTTCTAGCCCTTCGCAATCCCTPATATTCATCAACCAAC	180
QY	181	CCCTTGCCTCCCGATGTCAGCGCTGCGCGCGGGCTGGACCTGTTTCCGCAACCGGC	240
DB	181	CCCTTGCCTCCCGATGTCAGCGCTGCGCGCGGGCTGGACCTGTTTCCGCAACCGGC	240
QY	241	CGACCACTCGGCTCCGCTTGGCGTGACCGCGCCAGCGCCCGCTGCTCACTCGT	300
DB	241	CGACCACTCGGCTCCGCTTGGCGTGACCGCGCCAGCGCCCGCTGCTCACTCGT	300
QY	301	AGACCTACCAAGCTGGGGCGCGCTAACCGCGGTGCTCCGCGCCATGACACCCCG	360
DB	301	AGACCTACCAAGCTGGGGCGCGCTAACCGCGGTGCTCCGCGCCATGACACCCCG	360
QY	361	CGAGTGCCTGATGTGACCTCCCGCGGCGCATCTTGGCGCGGAGTATACCTATCAACA	420
DB	361	CGAGTGCCTGATGTGACCTCCCGCGGCGCATCTTGGCGCGGAGTATACCTATCAACA	420
QY	421	TCCTCCCTTACCTTCCGTTGGCGCACCGGCACTAACTGGTTCTTTATGCGGCCCTCT	480
DB	421	TCCTCCCTTACCTTCCGTTGGCGCACCGGCACTAACTGGTTCTTTATGCGGCCCTCT	480
QY	481	AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATAATGGCCACGGAGCTTCT	540
DB	481	AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATAATGGCCACGGAGCTTCT	540
QY	541	AATATGCGCAGTACCGGTTGGCGGTCACCAATCCGTTACCGCGCTGTTCCCAAT	600
DB	541	AATATGCGCAGTACCGGTTGGCGGTCACCAATCCGTTACCGCGCTGTTCCCAAT	600
QY	601	GCTGTGCGGGTTACGCACTCTCATCTCTTGGCCACAGACCAACACCCCGACG	660
DB	601	GCTGTGCGGGTTACGCACTCTCATCTCTTGGCCACAGACCAACACCCCGACG	660
QY	661	TCCGTTGATATGAATTCAATAAAGTTCGATATTTAGTTCAGCGCGGCATATA	720
DB	661	TCCGTTGATATGAATTCAATAAAGTTCGATATTTAGTTCAGCGCGGCATATA	720
QY	721	GCCTCTGAGCTTGATGCCAAGTGAGCGCTACACTATGTAACCAAGGCTGCGCTCC	780
DB	721	GCCTCTGAGCTTGATGCCAAGTGAGCGCTACACTATGTAACCAAGGCTGCGCTCC	780
QY	781	GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGCTTGTATGCTTTGCATA	840

DB	781	GTCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGCTTGTATGCTTTGCATA	840
QY	841	CATGCTCACTCGTAAATTCCTATATAACACCTATACCGGTGCGCTCGGCTGTTG	900
DB	841	CATGCTCACTCGTAAATTCCTATATAACACCTATACCGGTGCGCTCGGCTGTTG	900
QY	901	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACACCAATACGCGGTC	960
DB	901	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACACCAATACGCGGTC	960
QY	961	TCCCGTTATTCAGCACTGCTCGCAACCGCTTCGTCGCGGTGCGGACGCGGAG	1020
DB	961	TCCCGTTATTCAGCACTGCTCGCAACCGCTTCGTCGCGGTGCGGACGCGGAG	1020
QY	1021	CTCACCAACACCGCTTACCGCTTATGAGGAGCTCTATTTACTAGTACTAATGTT	1080
DB	1021	CTCACCAACACCGCTTACCGCTTATGAGGAGCTCTATTTACTAGTACTAATGTT	1080
QY	1081	GTCGCTGAGATCGGCGCGGATAGCCCTCACCCCTGTTCAACCTTCTGACACTGCTT	1140
DB	1081	GTCGCTGAGATCGGCGCGGATAGCCCTCACCCCTGTTCAACCTTCTGACACTGCTT	1140
QY	1141	GCGCGCTCGCAGCAAGATTGATTGCTCGGCTGCTGGCGAGCTGTTCTACTCCGTC	1200
DB	1141	GCGCGCTCGCAGCAAGATTGATTGCTCGGCTGCTGGCGAGCTGTTCTACTCCGTC	1200
QY	1201	GTTGCTCAGCAATAGGCGCGGATAGCTGTAAGTTGATATACATCTGAGAGATGCTCAG	1260
DB	1201	GTTGCTCAGCAATAGGCGCGGATAGCTGTAAGTTGATATACATCTGAGAGATGCTCAG	1260
QY	1261	CAGGATTAAGGTTATTCGAATCCGCTGACATGACCTCGGAGAACTCTGCTGGTTATT	1320
DB	1261	CAGGATTAAGGTTATTCGAATCCGCTGACATGACCTCGGAGAACTCTGCTGGTTATT	1320
QY	1321	CAGGATTAAGTAAACCAATGAAATCGGCGAGCGCTTCTCAGCCCATCGCGC	1380
DB	1321	CAGGATTAAGTAAACCAATGAAATCGGCGAGCGCTTCTCAGCCCATCGCGC	1380
QY	1381	CTTTCTCTGCTTCTCGAGCTAATGATGCTTGGCTCTCTCTCAGCTGCGGAGTAT	1440
DB	1381	CTTTCTCTGCTTCTCGAGCTAATGATGCTTGGCTCTCTCTCAGCTGCGGAGTAT	1440
QY	1441	GACAGTCCACTTATGGCTTCTCGACTGGCCCAAGTTATGTTTCTGACTGTCGACTG	1500
DB	1441	GACAGTCCACTTATGGCTTCTCGACTGGCCCAAGTTATGTTTCTGACTGTCGACTG	1500
QY	1501	GTTAATGTTGCGACCGCGCGGAGCGCTTGGCGGTCGCTCGATTTGGACCAAGTACA	1560
DB	1501	GTTAATGTTGCGACCGCGCGGAGCGCTTGGCGGTCGCTCGATTTGGACCAAGTACA	1560
QY	1561	CTTGAAGTCCGCGCTTCTCACCATTCAGCTGCTGAGAGACCTTCTTGTCTCGCG	1620
DB	1561	CTTGAAGTCCGCGCTTCTCACCATTCAGCTGCTGAGAGACCTTCTTGTCTCGCG	1620
QY	1621	CTCGCGGTAAGCTCTCTTTCTGAGGAGGAGGACCAACTAAAGCCGGTACCTTATAAT	1680
DB	1621	CTCGCGGTAAGCTCTCTTTCTGAGGAGGAGGACCAACTAAAGCCGGTACCTTATAAT	1680
QY	1681	TATAACACCACTGCTAGCGCAACTGCTTGTGAGAGTCCGCGCGGACCGGTCGCT	1740
DB	1681	TATAACACCACTGCTAGCGCAACTGCTTGTGAGAGTCCGCGCGGACCGGTCGCT	1740
QY	1741	ATTTCACCTTACACCTAGCTGGGTGCTGGTCCGCTCTCCATTTCTGCGGTTGCCGTT	1800
DB	1741	ATTTCACCTTACACCTAGCTGGGTGCTGGTCCGCTCTCCATTTCTGCGGTTGCCGTT	1800
QY	1801	TTAGCCCCCACTCTGCGCTAGCATGCTTGAAGATACCTTGAGTACCTCCGCGCGC	1860
DB	1801	TTAGCCCCCACTCTGCGCTAGCATGCTTGAAGATACCTTGAGTACCTCCGCGCGC	1860
QY	1861	CATACCTTTGATGATTTCTGCGCAGAGTGGCGCCCTCTGCTTCAAGGCTGCGCTTC	1920

Db 1861 CATACCTTTGATGATTTCTGTCCAGAGTGCCGCCCTTGGCGCTTCAGGGCTGCGCTTTC 1920
 QY 1921 CAGTCTACTGTGCTGAGCTTCAGCGCTTAAGATGAAGTGGAAGTAACTCGGAGTTG 1980
 Db 1921 CAGTCTACTGTGCTGAGCTTCAGCGCTTAAGATGAAGTGGAAGTAACTCGGAGTTG 1980
 QY 1981 TAG 1983
 Db 1981 TAG 1983

RESULT 2
 AAT27107
 ID AAT27107 standard; DNA; 2049 BP.
 AC AAT27107;
 XX 16-OCT-2003 (revised)
 DT 06-AUG-1996 (first entry)
 XX
 DE Hepatitis E virus (Burma strain) ORF-2.
 XX
 KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
 KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
 KW capsid; ds.
 XX
 OS Hepatitis E virus; Burma strain.
 XX
 PN WO9612807-A2.
 XX
 PD 02-MAY-1996.
 XX
 PF 23-OCT-1995; 95WO-US013703.
 XX
 PR 24-OCT-1994; 94US-00327952.
 PR 13-OCT-1995; 95US-00542634.
 XX
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 XX
 PI Fuerst TR, Mcatee CP, Yarbrough PO, Zhang Y;
 XX
 DR WPI; 1996-230608/23.
 DR P-PSDB; AAR96089.
 XX
 PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
 XX diagnostic reagents for determining HEV infection and in vaccines.
 ES Disclosure; Page 69-70; 125pp; English.
 XX
 CC The hepatitis E virus (HEV) Burma strain ORF-2 (AAT27107) codes for the
 CC putative capsid protein (AAR96089). PCR amplification of ORF-2 allows
 CC prodn. of capsid protein or of C-terminal regions of the capsid protein
 CC (see also AAR96091, AAR96093 and AAR96095) and expression in Spodoptera
 CC frugiperda Sf9 insect cells using a baculovirus vector provides
 CC recombinant C-terminal regions (see also AAR96101 and AAR96103) useful as
 CC diagnostic reagents and in vaccines. The HEV Mexico strain ORF-2
 CC (AAT27108) may similarly be used. (Updated on 16-OCT-2003 to standardise
 CC OS field)
 XX
 SQ Sequence 2049 BP; 344 A; 686 C; 496 G; 523 T; 0 U; 0 Other;

Query Match 100.0%; Score 1983; DB 2; Length 2049;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCTCGGCTATTTGTTGCTGCTCTCTATGTTTTGCTATGCTGCGCGGCA 60
 Db 1 ATGGCGCTCGGCTATTTGTTGCTGCTCTCTATGTTTTGCTATGCTGCGCGGCA 60
 QY 61 CCGCCCGTCTAGCGCTCTGGCGCGCGCTGCTGGCGCGCGAGCGGTTCCGGCGTGGT 120
 Db 61 CCGCCCGTCTAGCGCTCTGGCGCGCGCTGCTGGCGCGCGAGCGGTTCCGGCGTGGT 120

QY 121 TTCTGGGTGACCGGGTTGATTTCTCAGCCCTTCGCAATCCCTTATTTATTCATCAACAAC 180
 Db 121 TTCTGGGTGACCGGGTTGATTTCTCAGCCCTTCGCAATCCCTTATTTATTCATCAACAAC 180
 QY 181 CCCTTTCGCCCCGATGTACCGCTCGCGCGGGGTGAGACTCTGTTGCGCAACCCCGCC 240
 Db 181 CCCTTTCGCCCCGATGTACCGCTCGCGCGGGGTGAGACTCTGTTGCGCAACCCCGCC 240
 QY 241 CGACCACTCGGCTCGCTTGGGTGACGAGGCCGAGCGCCCGTTCGCTTCACGTCGT 300
 Db 241 CGACCACTCGGCTCGCTTGGGTGACGAGGCCGAGCGCCCGTTCGCTTCACGTCGT 300
 QY 301 AGACCTACACAGCTGGGCGCGGCTAACCGGSGTCTCCGCGCCCATGACACCCCG 360
 Db 301 AGACCTACACAGCTGGGCGCGGCTAACCGGSGTCTCCGCGCCCATGACACCCCG 360
 QY 361 CCAGTGCCTGATGTGCACTTCCCGCGCGGCATCTTTGCGCGGCGAGTATAAATCAACA 420
 Db 361 CCAGTGCCTGATGTGCACTTCCCGCGCGGCATCTTTGCGCGGCGAGTATAAATCAACA 420
 QY 421 TCTCCCTTACCTCTTCCGTGCGCAACCGGCTAACCTGTTCTTTATGCGCGCCCTCT 480
 Db 421 TCTCCCTTACCTCTTCCGTGCGCAACCGGCTAACCTGTTCTTTATGCGCGCCCTCT 480
 QY 481 AGTCCGCTTTTACCCCTTTCAGGAGCGCACCAATACCCATATAATGCGGCAAGCTTCT 540
 Db 481 AGTCCGCTTTTACCCCTTTCAGGAGCGCACCAATACCCATATAATGCGGCAAGCTTCT 540
 QY 541 AATTAATGCGGAGTACCGGGTTGCGCGTGCACCAATCCGTTACCGCGCGCTGCTCCCAAT 600
 Db 541 AATTAATGCGGAGTACCGGGTTGCGCGTGCACCAATCCGTTACCGCGCGCTGCTCCCAAT 600
 QY 601 GCTGTGGGGTTAGCGCATCTCAATCTCTGCGGCGACAGACCAACACCGCGAG 660
 Db 601 GCTGTGGGGTTAGCGCATCTCAATCTCTGCGGCGACAGACCAACACCGCGAG 660
 QY 661 TCCGTTGATATGAATTCATTAACCTCGAGGAGTTCGTTATTTAGTCAGCGCGGCATA 720
 Db 661 TCCGTTGATATGAATTCATTAACCTCGAGGAGTTCGTTATTTAGTCAGCGCGGCATA 720
 QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTTACATCTGTAACCAAGGCTGGCGCTCC 780
 Db 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTTACATCTGTAACCAAGGCTGGCGCTCC 780
 QY 781 GTGAGACTCTGGGGTGGCTGAGAGGAGCTACCTGCTGTTGTTATGCTTTGCTGATA 840
 Db 781 GTGAGACTCTGGGGTGGCTGAGAGGAGCTACCTGCTGTTGTTATGCTTTGCTGATA 840
 QY 841 CATGGCTCACCTGTAATTTCTTACTATAACCTATACCGCTATACCGGCTGGCGCTGTTG 900
 Db 841 CATGGCTCACCTGTAATTTCTTACTATAACCTATACCGGCTGGCGCTGTTGTTG 900
 QY 901 GACTTTCCTTTCAGCTTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTCC 960
 Db 901 GACTTTCCTTTCAGCTTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTCC 960
 QY 961 TCCGTTATTCAGCACTGCTGCGCACCGCTTCTGTCGGTGGGAGGAGCTGCGGAG 1020
 Db 961 TCCGTTATTCAGCACTGCTGCGCACCGCTTCTGTCGGTGGGAGGAGCTGCGGAG 1020
 QY 1021 CTCACCAACCGGCTGCTTACCGGCTTTATGAAGGACCTCTATTTTACTAGTAAATGCT 1080
 Db 1021 CTCACCAACCGGCTGCTTACCGGCTTTATGAAGGACCTCTATTTTACTAGTAAATGCT 1080
 QY 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTT 1140
 Db 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTT 1140
 QY 1141 GCGGCGCTGCGGACAGAAATTTGTCGGCTGGTGGCGAGCTGTTCTACTCCCGTCCC 1200
 Db 1141 GCGGCGCTGCGGACAGAAATTTGTCGGCTGGTGGCGAGCTGTTCTACTCCCGTCCC 1200
 QY 1201 GTTCTCTCAGCAATGGCGAGCCGACTGTTAAGTTGTATATCATCTGTAGAGAAATGCTCAG 1260

Db 1201 GTTGCTCTAGCAATGGGCGGCGAGCTGTTAGTTGTATATATCTGTAGAGATGCTCAG 1260
QY 1261 CAGGATAAGGGTATTGCAATCCGACATGACATTTGACCTCGGAGAAATCTCGTGTGGTTATT 1320
Db 1261 CAGGATAAGGGTATTGCAATCCGACATGACATTTGACCTCGGAGAAATCTCGTGTGGTTATT 1320
QY 1321 CAGGATAATGATAACCAATGAAACAGATCGGCGGACGCTTCTCCAGCCCCATCGCGC 1380
Db 1321 CAGGATAATGATAACCAATGAAACAGATCGGCGGACGCTTCTCCAGCCCCATCGCGC 1380
QY 1381 CTTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTACCGCTGCGGAGTAT 1440
Db 1381 CTTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTACCGCTGCGGAGTAT 1440
QY 1441 GACCACTCCACTTATGGCTCTTCCAGTGGCCCAAGTTATGTTTCTGACTCTGTGACCTTG 1500
Db 1441 GACCACTCCACTTATGGCTCTTCCAGTGGCCCAAGTTATGTTTCTGACTCTGTGACCTTG 1500
QY 1501 GTTAATGTTGCGACCGGCGGCGGCGGCTTGGCGGTCGCTCGATTGGACCAAGGTACA 1560
Db 1501 GTTAATGTTGCGACCGGCGGCGGCGGCTTGGCGGTCGCTCGATTGGACCAAGGTACA 1560
QY 1561 CTTGACGGTGGCGGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCCTGCGG 1620
Db 1561 CTTGACGGTGGCGGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCCTGCGG 1620
QY 1621 CTCGCGGGTAAGCTCTCTTTCTGGAGGCGAGGCAACAATAAGCGGGTACCTTATAAT 1680
Db 1621 CTCGCGGGTAAGCTCTCTTTCTGGAGGCGAGGCAACAATAAGCGGGTACCTTATAAT 1680
QY 1681 TATAACACCACTCTAGCAGCAACTGCTTGTGAGAAATGCGCGCGGACCGGTCGCT 1740
Db 1681 TATAACACCACTCTAGCAGCAACTGCTTGTGAGAAATGCGCGCGGACCGGTCGCT 1740
QY 1741 ATTTCACCTTACACCACTAGCCTGGGTGCTGGTCCCGTCTCCATTCTCGGTTGCGGTT 1800
Db 1741 ATTTCACCTTACACCACTAGCCTGGGTGCTGGTCCCGTCTCCATTCTCGGTTGCGGTT 1800
QY 1801 TTAGCCCCCCTCTGCTAGCCTTGTGAGATGCTTGTGAGATACCTTGTGACTACCTGCGCGGCC 1860
Db 1801 TTAGCCCCCCTCTGCTAGCCTTGTGAGATGCTTGTGAGATACCTTGTGACTACCTGCGCGGCC 1860
QY 1861 CATACCTTTGATGATTTCTGCGGAGTGCGCGGCGGCTTCCAGGCTTCCAGGCTTCCGTTTC 1920
Db 1861 CATACCTTTGATGATTTCTGCGGAGTGCGCGGCGGCTTCCAGGCTTCCAGGCTTCCGTTTC 1920
QY 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTTAAGATGAAGTGGGTAAACTCGGAGTTG 1980
Db 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTTAAGATGAAGTGGGTAAACTCGGAGTTG 1980
QY 1981 TAG 1983
Db 1981 TAG 1983

RESULT 3
AAQ47129
ID AAQ47129 standard; DNA; 2094 BP.

XX AC AAQ47129;
XX DT 24-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 11-JAN-1994 (first entry)
XX HEV third and second ORFs.

XX Enterically transmitted non-A non-B hepatitis; ET-NANB; vaccine; ss.
XX Hepatitis E virus; Burma strain.
XX OS
XX Key Location/Qualifiers

FT CDS 5. .376
FT /*tag= a
FT /label= ORF3
FT 46. .2094
FT /*tag= b
FT /label= ORF2
FT 275. .376
FT misc_RNA /*tag= c
FT /label= 406.4-2
FT 718. .2029
FT /*tag= d
FT misc_RNA /label= C2
FT 1045. .2029
FT /*tag= e
FT /label= SG3
FT 1882. .2029
FT /*tag= f
FT /label= 406.3-2
XX WO9314116-A1.
XX 22-JUL-1993.
XX 15-JAN-1993; 93WO-US0000459.
XX 17-JAN-1992; 92US-00822335.
XX 01-MAY-1992; 92US-00876941.
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX (USSH) US SEC DEPT HEALTH.
XX Reyes GR, Bradley DW, Tam AW, Carl M;
XX WPI; 1993-243144/30.
XX P-PSDB; AAR38784, AAR38785.
XX New immunogenic hepatitis E virus (HEV) peptide(s) - are from the ORF1,
XX ORF2 and ORF3 regions of HEV, useful as a vaccine against HEV infection.
XX Disclosure; Fig 7; 48pp; English.
XX Immunogenic hepatitis E virus (HEV) peptides are selected from the ORF1,
XX ORF2 and ORF3 regions of HEV. The peptides can be used in vaccines to
XX prevent infection by HEV. The antibodies can neutralise and block HEV
XX infection and can be used to prevent or treat HEV infection. The peptides
XX and antibodies can also be used as diagnostic reagents. (Updated on 25-
XX MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS
XX field)
SQ Sequence 2094 BP; 353 A; 697 C; 508 G; 536 T; 0 U; 0 Other;
Query Match 100.0%; Score 1983; DB 2; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCTGCGGCCA 60
Db 46 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCTGCGGCCA 105
QY 61 CCGCCCGGTCAGCGCTGTGCGCCCGCTGTGGCGGCGGAGCGGCTTCCGCGGTGGT 120
Db 106 CCGCCCGGTCAGCGCTGTGCGCCCGCTGTGGCGGCGGAGCGGCTTCCGCGGTGGT 165
QY 121 TTCTGGGGTGACCGGGTTGATTTCTAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 166 TTCTGGGGTGACCGGGTTGATTTCTAGCCCTTCGCAATCCCTATATTCATCAACCAAC 225
QY 181 CCCTTCCCGCCCGGATGTCACCGCTGCGGCGGCGGCTGAGACTCTGTGTCCCAACCCGCC 240
Db 226 CCCTTCCCGCCCGGATGTCACCGCTGCGGCGGCGGCTGAGACTCTGTGTCCCAACCCGCC 285
QY 241 CGACCACTCGGCTCCCGCTTGGCGTGACCGGCCCGGCGGCTGCTCAGCTCGT 300

286 CGACCACTCGGCTCGGCTTGGGTGACCAAGGCCAGCGCCCGGTCCTCAAGTCGT 345
QY
301 AGACCTACACAGCTGGGCGCGCGCTAAACGGGGTGGCTCCGGCCCATGACACCCCG 360
Db
346 AGACCTACACAGCTGGGCGCGCGCTAAACGGGGTGGCTCCGGCCCATGACACCCCG 405
QY
361 CCAGTGCCTGATGTGCACTCCGCGCGCGCCATCTTGGCGCGGAGTATAACCTATCAACA 420
Db
406 CCAGTGCCTGATGTGCACTCCGCGCGCGCCATCTTGGCGCGGAGTATAACCTATCAACA 465
QY
421 TCTCCGCTTACCTCTTCGTGGCGACCGGCACTAACCTGGTGTCTTTATGCGCGCCCTCTT 480
Db
466 TCTCCGCTTACCTCTTCGTGGCGACCGGCACTAACCTGGTGTCTTTATGCGCGCCCTCTT 525
QY
481 AGTCCGCTTTTACCCCTTCAGGACCGGCAACCAATACCATATAATGGCCACGGAAGTCTT 540
Db
526 AGTCCGCTTTTACCCCTTCAGGACCGGCAACCAATACCATATAATGGCCACGGAAGTCTT 585
QY
541 AATTATGCCAGTACCGGGTTGCCGTGCCCAATCCGTTACCGCCCGCTGGTCCCCCAAT 600
Db
586 AATTATGCCAGTACCGGGTTGCCGTGCCCAATCCGTTACCGCCCGCTGGTCCCCCAAT 645
QY
601 GCTGTGGGGTTAGCCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCGAGC 660
Db
646 GCTGTGGGGTTAGCCCATCTCCATCTCATTTCTGGCCACAGACCAACACCCGAGC 705
QY
661 TCCGTTGATATGAATTCATTAACCTCGACGGATGTTGTTATTTAGTCCAGCCCGGCATA 720
Db
706 TCCGTTGATATGAATTCATTAACCTCGACGGATGTTGTTATTTAGTCCAGCCCGGCATA 765
QY
721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCCCTACACTATCGTAACCAAGGCTGGGGCTCC 780
Db
766 GCCTCTGAGCTTGTGATCCCAAGTGAAGCCCTACACTATCGTAACCAAGGCTGGGGCTCC 825
QY
781 GTCGAGACCTCTGGGGTGGCTGAGAGAGGCTACTCTGGTCTTGTATGCTTTGCGATA 840
Db
826 GTCGAGACCTCTGGGGTGGCTGAGAGAGGCTACTCTGGTCTTGTATGCTTTGCGATA 885
QY
841 CATGCTCACTCGTAAATTCCTATATAATACACCTATACCGGTAAACCAATACGGGGTGG 900
Db
886 CATGCTCACTCGTAAATTCCTATATAATACACCTATACCGGTAAACCAATACGGGGTGG 945
QY
901 GACTTTGGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTGG 960
Db
946 GACTTTGGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTGG 1005
QY
961 TCCCGTTATTCAGCACTGCTGCCACCGCTTCGTGGGGTGGGAGCGGAGTCCCGAG 1020
Db
1006 TCCCGTTATTCAGCACTGCTGCCACCGCTTCGTGGGGTGGGAGCGGAGTCCCGAG 1065
QY
1021 CTCACCAACCGGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGCT 1080
Db
1066 CTCACCAACCGGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGCT 1125
QY
1081 GTCGGTGAAGTGGCGCGGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1140
Db
1126 GTCGGTGAAGTGGCGCGGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1185
QY
1141 GCGGCTTCCGACAGAAATTTGCTGGGTGGTGGCGAGCTGTTCTACCTCCCGTCC 1200
Db
1186 GCGGCTTCCGACAGAAATTTGCTGGGTGGTGGCGAGCTGTTCTACCTCCCGTCC 1245
QY
1201 GTTGTCTCAGCAATGGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAACTCTCAG 1260
Db
1246 GTTGTCTCAGCAATGGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAACTCTCAG 1305
QY
1261 CAGGATAAGGTTATTTGCAATCCCGCATGACATTGACCTCGGAGAACTCTCGTGGTTATT 1320
Db
1306 CAGGATAAGGTTATTTGCAATCCCGCATGACATTGACCTCGGAGAACTCTCGTGGTTATT 1365
QY
1321 CAGGATTATGATAACCAATGAAACAAGATCGGCGGAGCCCTTCTCCAGCCCATCGCGC 1380
Db
1366 CAGGATTATGATAACCAATGAAACAAGATCGGCGGAGCCCTTCTCCAGCCCATCGCGC 1425

QY 1381 CTTTTCTCTGCTCTTCGAGCTAATGATGCTTTTGGCTCTCTCTCACCGCTGCCGAGTAT 1440
Db 1426 CTTTTCTCTGCTCTTCGAGCTAATGATGCTTTTGGCTCTCTCTCACCGCTGCCGAGTAT 1485
QY 1441 GACCACTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 1486 GACCACTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1545
QY 1501 GTTAATGTTGACCGGCGCGAGCGGCTGGTCCCGTCTCGATTGGAACCAAGGTCA 1560
Db 1546 GTTAATGTTGACCGGCGCGAGCGGCTGGTCCCGTCTCGATTGGAACCAAGGTCA 1605
QY 1561 CTTGACGGTGGCCCTCTCCACCACTCCAGCAGTACTCGAAGACCTTCTTTCTGCTGGCG 1620
Db 1606 CTTGACGGTGGCCCTCTCCACCACTCCAGCAGTACTCGAAGACCTTCTTTCTGCTGGCG 1665
QY 1621 CTTCCGGTAACTCTCTTTCTGGAGGAGGAGCAACAACCTAAAGCCGGGTACCTTATAT 1680
Db 1666 CTTCCGGTAACTCTCTTTCTGGAGGAGGAGCAACAACCTAAAGCCGGGTACCTTATAT 1725
QY 1681 TATAACCACTGCTAGCAACCACTGCTTGTGAGAAATGCCCGGAGCACCGGTCGT 1740
Db 1726 TATAACCACTGCTAGCAACCACTGCTTGTGAGAAATGCCCGGAGCACCGGTCGT 1785
QY 1741 ATTTCACCTTACCACTAGCTGCTGCTGCTCCGCTCTCCATTTCTGCGGTTGCCGTT 1800
Db 1786 ATTTCACCTTACCACTAGCTGCTGCTGCTCCGCTCTCCATTTCTGCGGTTGCCGTT 1845
QY 1801 TTAGCCCCCACTCTGCGCTAGCACTTGTGAGGATACCTTGGACTACCTGCCCGCGC 1860
Db 1846 TTAGCCCCCACTCTGCGCTAGCACTTGTGAGGATACCTTGGACTACCTGCCCGCGC 1905
QY 1861 CATACCTTTGATGATTTCTGCCAGAGTCCGCCCTTTGGCTTCAGGGCTGGCGTTTC 1920
Db 1906 CATACCTTTGATGATTTCTGCCAGAGTCCGCCCTTTGGCTTCAGGGCTGGCGTTTC 1965
QY 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGATGAAGGTGGTAAACCTCGGAGTTG 1980
Db 1966 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGATGAAGGTGGTAAACCTCGGAGTTG 2025
QY 1981 TAG 1983
Db 2026 TAG 2028
RESULT 4
AAQ46813
ID AAQ46813 standard; DNA; 2094 BP.
XX AC AAQ46813;
XX 25-MAR-2003 (revised)
DT 14-FEB-1994 (first entry)
XX
XX Burma strain HEV ORF2 and ORF3.
KW Hepatitis E virus; vaccine; neutralising antibodies; infection; block;
open reading frame; antibodies; ss.
XX
XX Hepatitis E virus.
XX Key Location/Qualifiers
FT CDS 5..311
FT /*tag= a
FT /note= "ORF2"
FT CDS 46..2094
FT /*tag= b
FT /note= "ORF3"
XX
PN WO9314208-A2.
XX
PD 22-JUL-1993.

XX 19-JAN-1993; 93WO-US000475.
XX 17-JAN-1992; 92US-00822335.
PR 20-APR-1992; 92US-00870985.
XX
XX (GENE-) GENELABS TECHNOLOGIES INC.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Reyes GR, Bradley DW, Twu J, Purdy MA, Tam AM, Krawczynski KZ;
PI P-PSDB; AAR39305, AAR39306.
XX WPI; 1993-243223/30.
DR P-PSDB; AAR39305, AAR39306.
XX
XX Antigen and antibody vaccines against hepatitis B virus infection -
PT contain peptide(s) derived from capsid protein C-terminal or antibodies
PT against protein.
XX
XX Disclosure; Fig 7; 43pp; English.
XX
XX The sequence is that encoding Burma strain hepatitis B virus (HEV) open
CC reading frames ORF2 and ORF3. ORF2 encodes the putative capsid protein of
CC the virus, this protein or peptide fragments of it may be used in a
CC vaccine composition for immunising an individual against HEV. Antibodies
CC raised against these peptides can also be used in such vaccines. (Updated
CC on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 2094 BP; 353 A; 697 C; 508 G; 536 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1983; DB 2; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTGGGCTATTGTTGGTCTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
DB 46 ATGCGCCCTGGGCTATTGTTGGTCTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA 105
QY 61 CCGCCCGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGTG 120
DB 106 CCGCCCGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCG 165
QY 121 TTCTGGGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCG 180
DB 166 TTCTGGGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCG 225
QY 181 CCCTTGGCCCGGATGTCAGCGTGGCGGTGACCGGTGACCGGTGACCGGTGACCGGTG 240
DB 226 CCCTTGGCCCGGATGTCAGCGTGGCGGTGACCGGTGACCGGTGACCGGTGACCGGTG 285
QY 241 CGACCACTGGGTCCGCTTGGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCG 300
DB 286 CGACCACTGGGTCCGCTTGGGTGACCGGTGACCGGTGACCGGTGACCGGTGACCG 345
QY 301 AGACCTACCAAGCTGGGGCGCGCGCTGACCGGTGACCGGTGACCGGTGACCGGTG 360
DB 346 AGACCTACCAAGCTGGGGCGCGCGCTGACCGGTGACCGGTGACCGGTGACCGGTG 405
QY 361 CGAGTGGCTGATGTCAGCTCCGCGGCGCATCTTGGCGGCGGAGTATACCTATCAACA 420
DB 406 CGAGTGGCTGATGTCAGCTCCGCGGCGCATCTTGGCGGCGGAGTATACCTATCAACA 465
QY 421 TCTCCCTTACCTCTTCCGTGGCCACCGGCACTAACTGGTCTTATGCGCGCCCTCT 480
DB 466 TCTCCCTTACCTCTTCCGTGGCCACCGGCACTAACTGGTCTTATGCGCGCCCTCT 525
QY 481 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATATATGCGCACCGAGCTTCT 540
DB 526 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATATATGCGCACCGAGCTTCT 585
QY 541 AATATGCGCAGTACCGGGTGGCGGTGCGGACCAATCGTTACCGCGCGGTGTCGCCAAT 600
DB 586 AATATGCGCAGTACCGGGTGGCGGTGCGGACCAATCGTTACCGCGCGGTGTCGCCAAT 645

QY 601 GCTGTGCGGGTTAAGCCATCTCCATCTCATCTGCGCCACAGACCAACCAACCCCGACG 660
DB 646 GCTGTGCGGGTTAAGCCATCTCCATCTCATCTGCGCCACAGACCAACCAACCCCGACG 705
QY 661 TCCGTTGATATGAATTCATAAACCCTCGACGGATGTTCTGATATTTTATGTCAGCCG 720
DB 706 TCCGTTGATATGAATTCATAAACCCTCGACGGATGTTCTGATATTTTATGTCAGCCG 765
QY 721 GCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGCTGCGCTCC 780
DB 766 GCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGCTGCGCTCC 825
QY 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
DB 826 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 885
QY 841 CATGGCTCACTCGTAAATTCCTATATAACCCCTATACCGGTGCGGTGCGGTGTTG 900
DB 886 CATGGCTCACTCGTAAATTCCTATATAACCCCTATACCGGTGCGGTGCGGTGTTG 945
QY 901 GACTTTGCGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 960
DB 946 GACTTTGCGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 1005
QY 961 TCCGCTTATTCAGCACTGCTGCCACCGCTTGGTGGCGGTGCGGACGCGGACGCGGAG 1020
DB 1006 TCCGCTTATTCAGCACTGCTGCCACCGCTTGGTGGCGGTGCGGACGCGGACGCGGAG 1065
QY 1021 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTTATTTTACTAGTACTAATGCT 1080
DB 1066 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTTATTTTACTAGTACTAATGCT 1125
QY 1081 GTCGCTGAGTGGCGCGGAGTAGCCCTCACCTGTTCAACCTGCTGACACTCTGCTT 1140
DB 1126 GTCGCTGAGTGGCGCGGAGTAGCCCTCACCTGTTCAACCTGCTGACACTCTGCTT 1185
QY 1141 GCGCGCTGCGGACAGAAATGATTTGCTGGCTGGTGGCGGAGCTTCTACTCCCGTCCC 1200
DB 1186 GCGCGCTGCGGACAGAAATGATTTGCTGGCTGGTGGCGGAGCTTCTACTCCCGTCCC 1245
QY 1201 GTTGTCTGAGCCAAATGCGGAGCGGACTGTTAAGTTGTATACATCTCTAGAGAAATGCT 1260
DB 1246 GTTGTCTGAGCCAAATGCGGAGCGGACTGTTAAGTTGTATACATCTCTAGAGAAATGCT 1305
QY 1261 CAGGATAAGGGTATTCGAATCCCGCATGACATTCGCGGAGAACTCTGTGCTGTTATT 1320
DB 1306 CAGGATAAGGGTATTCGAATCCCGCATGACATTCGCGGAGAACTCTGTGCTGTTATT 1365
QY 1321 CAGGATTATGATAACCAACATGAACAGATCGGCGGAGCTTCTCAGCCCCCATCGCGC 1380
DB 1366 CAGGATTATGATAACCAACATGAACAGATCGGCGGAGCTTCTCAGCCCCCATCGCGC 1425
QY 1381 CCTTTCTGCTTCTGAGCTAATGATGCTGTTTGGCTCTCTCTCAGCGTGGCGAGTAT 1440
DB 1426 CCTTTCTGCTTCTGAGCTAATGATGCTGTTTGGCTCTCTCTCAGCGTGGCGAGTAT 1485
QY 1441 GACCACTGCTTATGCTTCTGCTGCGGCGGAGTATGTTTCTGACTCTGTGACCTTG 1500
DB 1486 GACCACTGCTTATGCTTCTGCTGCGGCGGAGTATGTTTCTGACTCTGTGACCTTG 1545
QY 1501 GTTAAATGTTGCGACCGCGCGGCTTGGCGGCTGCGTTCGATTGGACCAAGGTCA 1560
DB 1546 GTTAAATGTTGCGACCGCGCGGCTTGGCGGCTGCGTTCGATTGGACCAAGGTCA 1605
QY 1561 CTTGAGCTGCGGCGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCGG 1620
DB 1606 CTTGAGCTGCGGCGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCGG 1665
QY 1621 CTCGCGGTGAAGCTCTCTTTCTGGGAGGAGGACCAACATAAGCCGGGTACCTTATAAT 1680
DB 1666 CTCGCGGTGAAGCTCTCTTTCTGGGAGGAGGACCAACATAAGCCGGGTACCTTATAAT 1725
QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTTCGAGAAATCCGCGCGGCGACCGGGT 1740

```

Db 1726 TATAACACCACTGCTAGCGACCACTGCTTGTGAGAAATGCGCGGACCGGGTCGCT 1785
QY 1741 ATTTCCACTTACACCACTAGCCTGGGTGCTGCTCCGCTCTCCATTTCTGCGGTTCGGTT 1800
Db 1786 ATTTCCACTTACACCACTAGCCTGGGTGCTGCTCCGCTCTCCATTTCTGCGGTTCGGTT 1845
QY 1801 TTAGCCCCCACTCTCGCTAGCATTTGCTTGGAGATACCTTGGACTACCTTGCCCGCGC 1860
Db 1846 TTAGCCCCCACTCTCGCTAGCATTTGCTTGGAGATACCTTGGACTACCTTGCCCGCGC 1905
QY 1861 CATACCTTTGATGATTTCTGCCAGAGTGCCTGCGCTTGGAGTGCCTTGGAGTGCCTTTC 1920
Db 1906 CATACCTTTGATGATTTCTGCCAGAGTGCCTGCGCTTGGAGTGCCTTGGAGTGCCTTTC 1965
QY 1921 CAGTCTACTGCTGCTGAGCTTACGCGCTTACAGATGAAGTGGTAAACTCGGAGTTG 1980
Db 1966 CAGTCTACTGCTGCTGAGCTTACGCGCTTACAGATGAAGTGGTAAACTCGGAGTTG 2025
QY 1981 TAG 1983
Db 2026 TAG 2028

RESULT 5
AAV66321
ID AAV66321 standard; DNA; 7195 BP.
XX AC AAV66321;
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 23-DEC-1998 (first entry)
XX ET-NANB (HEV) Burma strain DNA sequence.
XX Enterically transmitted non A non B hepatitis virus; ET-NANB;
KW Hepatitis E virus; HEV; Burma HEV isolate; vaccine; diagnostic probe; ds.
XX Non-A.
XX non-B hepatitis virus.
XX Key Location/Qualifiers
FT CDS 28..5109
FT FT /*tag= a
FT /*note= " encodes AAW80196"
FT CDS 5106..5477
FT FT /*tag= c
FT /*note= " encodes AAW80198"
FT CDS 5147..7129
FT FT /*tag= b
FT /*note= " encodes AAW80197"
XX US5824649-A.
XX 20-OCT-1998.
XX 07-JUN-1995; 95US-00475807.
XX 17-JUN-1988; 88US-00208997.
XX 11-APR-1989; 89US-00336672.
XX 16-JUN-1989; 89US-00367486.
XX 13-OCT-1989; 89US-00420921.
XX 05-APR-1990; 90US-00505888.
XX 25-JUL-1994; 94US-00279823.
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX Bradley DW, Reyes GR, Krawczynski KZ, Tam A, Fry KE, Yarbrough PO;
XX WPI: 1998-582599/49.
XX P-PSDB; AAW80196, AAW80197, AAW80198.
XX

```

```

PT Hepatitis E virus proteins - useful for diagnosis or vaccine production
PT the virus.
XX Claim 1; Col 51-58; 47pp; English.
XX The present sequence represents the genome of the Burma strain of
CC enterically transmitted non A non B hepatitis virus (ET-NANB) (hepatitis
CC E virus (HEV)). The specification describes an isolated protein which is
CC specifically immunoreactive with antibodies present in individuals
CC infected with HEV and encoded by a sequence contained in an open reading
CC frame (ORF) of an HEV genome. The genome has a sequence that is more than
CC 70% identical to the ORF1 sequence from Burma HEV isolate. The protein is
CC used as a vaccine and a diagnostic probe for ET-NANB. (Updated on 25-MAR-
CC 2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 7195 BP; 1221 A; 2312 C; 1868 G; 1794 T; 0 U; 0 Other;
Query Match 100.0%; Score 1983; DB 2; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCTATTTTGTGTGCTCTCTCAATTTTTCCTTANGTGGCGGCCA 60
Db 5147 ATGCGCCCTCGGCTATTTTGTGTGCTCTCTCAATTTTTCCTTANGTGGCGGCCA 5206
QY 61 CGCGCCGCTCAGCGCTGCGCGCGCTGCTGGCGCGCGCGCGCGGTTCCGCGGTGT 120
Db 5207 CGCGCCGCTCAGCGCTGCGCGCGCTGCTGGCGCGCGCGCGGTTCCGCGGTGT 5266
QY 121 TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 5267 TTCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 5326
QY 181 CCCTTCGCCCCCGATGTACCGCTGCGCGCGCGGCTGACCTCTGTTTCGCAACCCCGC 240
Db 5327 CCCTTCGCCCCCGATGTACCGCTGCGCGCGGCTGACCTCTGTTTCGCAACCCCGC 5386
QY 241 CGACCACTCGGCTCGCTTGGGTGACCGAGCGCGCGCGCGCGGTTCCGCTCAGTGT 300
Db 5387 CGACCACTCGGCTCGCTTGGGTGACCGAGCGCGCGCGCGGTTCCGCTCAGTGT 5446
QY 301 AGACCTACACAGCTGGGGCGCGCGCTAACCGCGGTGCTCCGCGCCCATGACACCCCG 360
Db 5447 AGACCTACACAGCTGGGGCGCGCGCTAACCGCGGTGCTCCGCGCCCATGACACCCCG 5506
QY 361 CCAGTGCCTGATGTGCACTCCCGCGCGCGCATTTTGGCGCGCGAGTATTAACCTATCAACA 420
Db 5507 CCAGTGCCTGATGTGCACTCCCGCGCGCGCATTTTGGCGCGCGAGTATTAACCTATCAACA 5566
QY 421 TCTCCCTTACCTCTTCCGTGCGCACCGGCTACCTGCTTCTTTATGCGCGCGCTCTT 480
Db 5567 TCTCCCTTACCTCTTCCGTGCGCACCGGCTACCTGCTTCTTTATGCGCGCGCTCTT 5626
QY 481 AGTCCGCTTTTACCCCTTCAGGACCGGACCAATACCATATTAATGGCCACGAAAGTCT 540
Db 5627 AGTCCGCTTTTACCCCTTCAGGACCGGACCAATACCATATTAATGGCCACGAAAGTCT 5686
QY 541 AATATGCCAGTACCGGGTTCGCGTCCCAATTCGTTTACCGCGCGCTGTCGCCAAT 600
Db 5687 AATATGCCAGTACCGGGTTCGCGTCCCAATTCGTTTACCGCGCGCTGTCGCCAAT 5746
QY 601 GCTGTGGCGGTTCGCGCTTCCATCTCTCATTTCTGGCCACAGACCAACACCGCGGAG 660
Db 5747 GCTGTGGCGGTTCGCGCTTCCATCTCTCATTTCTGGCCACAGACCAACACCGCGGAG 5806
QY 661 TCCGTGTATATCAATTCATTAACCTTCGACGATGTTTCGTATTTTAGTCAGCCCGGATA 720
Db 5807 TCCGTGTATATCAATTCATTAACCTTCGACGATGTTTCGTATTTTAGTCAGCCCGGATA 5866
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAAAGGCTGGCGTCC 780
Db 5867 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAAAGGCTGGCGTCC 5926

```

QY 781 GTCGAGACCTCTGGGGTGGCTCAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
 DB 5927 GTCGAGACCTCTGGGGTGGCTCAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 5986
 QY 841 CATGGCTCACTCGTAATTCCTATATCAATACACCCCTATACCGGTGCCCTCGGGCTGTTG 900
 DB 5987 CATGGCTCACTCGTAATTCCTATATCAATACACCCCTATACCGGTGCCCTCGGGCTGTTG 6046
 QY 901 GACTTTGCCCTTTGAGCTTTGAGTTTCGAACTTTACCCCGGTAAACCAATACGCGGTC 960
 DB 6047 GACTTTGCCCTTTGAGCTTTGAGTTTCGAACTTTACCCCGGTAAACCAATACGCGGTC 6106
 QY 961 TCCGGTTATTCAGCACTGCTGCCACCGCTTCGTGCGGTGCGGAGCGGACTGCGAG 1020
 DB 6107 TCCGGTTATTCAGCACTGCTGCCACCGCTTCGTGCGGTGCGGAGCGGACTGCGAG 6166
 QY 1021 CTCACCAACAGGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
 DB 6167 CTCACCAACAGGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 6226
 QY 1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCTGTTCAACCTTGTCGACACTCTGCTT 1140
 DB 6227 GTCGGTGAGATCGGCGCGGATAGCCCTCACCTGTTCAACCTTGTCGACACTCTGCTT 6286
 QY 1141 GCGGGCTCGCGACAGATTTGTTGTCGGCTGCTGGCCAGCTGTTCTACTCCCGTCCC 1200
 DB 6287 GCGGGCTCGCGACAGATTTGTTGTCGGCTGCTGGCCAGCTGTTCTACTCCCGTCCC 6346
 QY 1201 GTTGTCTCAGCAATGGCGAGCGACTGTTAAAGTTGTATATCTGTAGAGAATGCTCAG 1260
 DB 6347 GTTGTCTCAGCAATGGCGAGCGACTGTTAAAGTTGTATATCTGTAGAGAATGCTCAG 6406
 QY 1261 CAGGATAAGGGTATTTGCAATCCGCGATGACATTTGACCTCGGAGAACTCTGTGTTATT 1320
 DB 6407 CAGGATAAGGGTATTTGCAATCCGCGATGACATTTGACCTCGGAGAACTCTGTGTTATT 6466
 QY 1321 CAGGATTATGATACCAACATGACAGATCGCGGAGCGCTTCTCCAGCCCATCGCG 1380
 DB 6467 CAGGATTATGATACCAACATGACAGATCGCGGAGCGCTTCTCCAGCCCATCGCG 6526
 QY 1381 CCTTTCTCTCTCTCAGCTAATGATGCTTTGGCTCTCTCTCAGCGTCCGAGTAT 1440
 DB 6527 CCTTTCTCTCTCTCAGCTAATGATGCTTTGGCTCTCTCTCAGCGTCCGAGTAT 6586
 QY 1441 GACAGTCCACTATAGCTCTTCGACCTGCGCCAGTTTATGTTCTGACTCTGTGACCTTG 1500
 DB 6587 GACAGTCCACTATAGCTCTTCGACCTGCGCCAGTTTATGTTCTGACTCTGTGACCTTG 6646
 QY 1501 GTTAATGTTGCGAGCGCGCGAGCGGCTTCCCGGTCGCTCGATTGCAACCAAGGTCACA 1560
 DB 6647 GTTAATGTTGCGAGCGCGCGAGCGGCTTCCCGGTCGCTCGATTGCAACCAAGGTCACA 6706
 QY 1561 CTTGACGGTCCGCCCTCTCCACCATCAGCAGTACTCGAAGACCTTTCTTGTCTCGCG 1620
 DB 6707 CTTGACGGTCCGCCCTCTCCACCATCAGCAGTACTCGAAGACCTTTCTTGTCTCGCG 6766
 QY 1621 CTCGGGTAAGCTCTCTTCTGGGAGCGGCAACCTAAGCGGTAACCTTATAT 1680
 DB 6767 CTCGGGTAAGCTCTCTTCTGGGAGCGGCAACCTAAGCGGTAACCTTATAT 6826
 QY 1681 TATAACACCACTGCTAGCGACCACTGCTTGTTCGAGAAATCGCGCGGCAACCGGTCGCT 1740
 DB 6827 TATAACACCACTGCTAGCGACCACTGCTTGTTCGAGAAATCGCGCGGCAACCGGTCGCT 6886
 QY 1741 ATTTCCACTTACCACTAGCCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTTGCGGTT 1800
 DB 6887 ATTTCCACTTACCACTAGCCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTTGCGGTT 6946
 QY 1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTTCAGATACCTTGGACTACCTTCCCGGCC 1860
 DB 6947 TTAGCCCCCACTCTGCGCTAGCAATGCTTTCAGATACCTTGGACTACCTTCCCGGCC 7006
 QY 1861 CATACTTTTGATGATTTCTTGCCCAAGAGTCCGCCGCCCTTGCCCTTCAGGGCTGCGCTTC 1920

DB 7007 CATACTTTTGATGATTTCTGCCAGAGTGCCGCCCTTTGGCTTCAGGCTGCGCTTTC 7066
 QY 1921 CAGTCTACTGTCGCTGAGCTTCAGGCTTAAGATGAAGGTGGTAAACTCGGAGTTG 1980
 DB 7067 CAGTCTACTGTCGCTGAGCTTCAGGCTTAAGATGAAGGTGGTAAACTCGGAGTTG 7126
 QY 1981 TAG 1983
 DB 7127 TAG 7129
 RESULT 6
 AAV54729
 ID AAV54729 standard; DNA; 7195 BP.
 XX AAV54729;
 XX AC
 XX 25-MAR-2003 (revised)
 DT 30-OCT-1998 (first entry)
 XX
 XX DNA sequence of the Burmese isolate of ET-NANB.
 XX
 XX Externically transmitted nonA/nonB hepatitis virus; identification; HEV;
 KW ET-NANB; detection; vaccine; ds.
 XX
 XX Hepatitis virus.
 XX
 XX Key Location/Qualifiers
 FH 28..5109
 FT CDS /tag= a
 FT /note= "see AAW71209"
 FT 5106..5477
 FT /tag= c
 FT /note= "see AAW71211"
 FT 5147..7129
 FT /tag= b
 FT /note= "see AAW71210"
 XX
 XX US5789559-A.
 XX 04-AUG-1998.
 XX 25-JUL-1994; 94US-00279823.
 XX 17-JUN-1988; 88US-00208997.
 PR 11-APR-1989; 89US-00336672.
 PR 19-JUN-1989; 89US-00367486.
 PR 13-OCT-1989; 89US-00420921.
 PR 05-APR-1990; 90US-00505888.
 PR 05-APR-1991; 91US-00681078.
 XX
 XX (GENE-) GENELABS TECHNOLOGIES INC.
 XX
 XX Bradley DW, Reyes GR, Krawczynski KZ, Tam A, Fry KE, Yarbough PO;
 XX WPI: 1998-446186/38.
 XX P-PSDB: AAW71209, AAW71210, AAW71211.
 DR
 DR Hepatitis E virus DNA - useful for e.g. virus detection and viral protein
 PT production.
 XX
 XX Disclosure; Col 49-56; 45pp; English.
 XX
 XX The present sequence represents the sequence of the Burmese isolate of an
 CC enterically transmitted nonA/nonB viral hepatitis agent (ET-NANB). The
 CC nucleic acid sequence may be used for identifying and sequencing the
 CC entire viral agent (also referred to as HEV), detecting ET-NANB in
 CC infected samples, e.g. by specific amplification of virus-derived DNA
 CC sequences and for producing recombinant viral proteins for use in
 CC vaccines. (Updated on 25-MAR-2003 to correct PF field.)
 XX
 XX Sequence 7195 BP; 1221 A; 2312 C; 1868 G; 1794 T; 0 U; 0 Other;

[illegible]

D b	6287	GGGGCGCTCCGACAGAAATTGATTTCTGTCGGCTGGTGCGCCAGCTGTTCTCTACTCCCGTCCC	6346
Q y	1201	GTGTCTCAGCCAAATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG	1260
D b	6347	GTGTCTCAGCCAAATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG	6406
Q y	1261	CAGGATAGGGTATTGCAATCCCGCATGACATTGACCTCGAGAAATCTCTGTGTGGTTATT	1320
D b	6407	CAGGATAGGGTATTGCAATCCCGCATGACATTGACCTCGAGAAATCTCTGTGTGGTTATT	6466
Q y	1321	CAGGATTATGATAACCAACATGAACAAGATCGCGCAGCGCTTCTCGAGCCCCATCGCGC	1380
D b	6467	CAGGATTATGATAACCAACATGAACAAGATCGCGCAGCGCTTCTCGAGCCCCATCGCGC	6526
Q y	1381	CTTTCTCTGTCTTCGAGCTAAATGATGTCTTTGGTCTCTCTCACCGTGCAGAGTAT	1440
D b	6527	CTTTCTCTGTCTTCGAGCTAAATGATGTCTTTGGTCTCTCTCACCGTGCAGAGTAT	6586
Q y	1441	GACCAGTCCACTTATGGCTCTTCGACGTGCCCGAGTTTATGTTTCTGTGACCTTG	1500
D b	6587	GACCAGTCCACTTATGGCTCTTCGACGTGCCCGAGTTTATGTTTCTGTGACCTTG	6646
Q y	1501	GTTAATGTTGCGACGGCGCGCAGGCGGTTCGCCGGTGGTTCGATTGGAACCAAGGTCA	1560
D b	6647	GTTAATGTTGCGACGGCGCGCAGGCGGTTCGCCGGTGGTTCGATTGGAACCAAGGTCA	6706
Q y	1561	CTTCACGGTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTCGCG	1620
D b	6707	CTTCACGGTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTCGCG	6766
Q y	1621	CTCCGGGTAAAGCTCTCTTTCTGGGAGCGAGGCACAACTAAAGCCGGTACCTTATAAT	1680
D b	6767	CTCCGGGTAAAGCTCTCTTTCTGGGAGCGAGGCACAACTAAAGCCGGTACCTTATAAT	6826
Q y	1681	TATTAACACACTGCTAGCGACCAACTGCTTGTGAGAAATCCGCGCGGCACCCGGTGCCT	1740
D b	6827	TATTAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATCCGCGCGGCACCCGGTGCCT	6886
Q y	1741	ATTTCCTACTTACCACTAGCCTGGGTGCTGGTCCCGTCTCCATTTCTCGGGTTCGGGTT	1800
D b	6887	ATTTCCTACTTACCACTAGCCTGGGTGCTGGTCCCGTCTCCATTTCTCGGGTTCGGGTT	6946
Q y	1801	TTAGCCCCCACTCTGGGCTAGCAATTGCTTGAGATACCTTGACTACCTCGCCCGCGCC	1860
D b	6947	TTAGCCCCCACTCTGGGCTAGCAATTGCTTGAGATACCTTGACTACCTCGCCCGCGCC	7006
Q y	1861	CATACCTTTTGATGATTTCTGCCAGAGTGGCGGCCCTTGGCCTTCAGGGCTCGCTTTC	1920
D b	7007	CATACCTTTTGATGATTTCTGCCAGAGTGGCGGCCCTTGGCCTTCAGGGCTCGCTTTC	7066
Q y	1921	CAGTCTACTGTGCTGAGCTTCAGCGCTTAAAGATGAAGTGGGTAAACTCGGAGTTG	1980
D b	7067	CAGTCTACTGTGCTGAGCTTCAGCGCTTAAAGATGAAGTGGGTAAACTCGGAGTTG	7126
Q y	1981	TAG 1983	
D b	7127	TAG 7129	

RESULTS

RESULT 8
AAFB3495

AAF83493
ID AAF83495 standard: DNA: 7195 BP.

XX
CSC07977

AC AAF83495;

XX
XX

DT 23-JUL-2001 (first entry)

5
 4
 3
 2
 1
 0

DE ET-NANB V

KW HEV; ente

XX	Key	Location/Qualifiers
FH	CDS	28..5109
FT		/*tag= a
FT		/product= "ORF1"
FT	CDS	5106..5477
FT		/*tag= c
FT		/product= "ORF3"
FT	CDS	5147..7129
FT		/*tag= b
FT		/product= "ORF2"
XX		
PN	US229005-B1.	
XX		
XX	08-MAY-2001.	
XX		
PF	03-AUG-1998;	98US-00128275.
XX		
PR	17-JUN-1988;	88US-00208997.
PR	11-APR-1989;	89US-00336672.
PR	16-JUN-1989;	89US-00367486.
PR	13-OCT-1989;	89US-00420921.
PR	05-APR-1990;	90US-00505888.
PR	05-APR-1991;	91US-00681078.
PR	25-JUL-1994;	94US-00279823.
XX		
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PA	(GENE-) GENELABS TECHNOLOGIES INC.	
XX		
PI	Reyes GR, Yarbrough PO, Bradley DW, Krawczynski KZ, Tam AW;	
PI	Fry KE;	
XX		
DR	WPI; 2001-342705/36.	
DR	P-PSDB; AAB62522, AAB62523, AAB62524.	
XX		
PT	New DNA sequences of enterically transmitted non-A/non-B (ET-NANB)	
PT	hepatitis viral agent, useful in diagnosing infection by an enterically	
PT	transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.	
XX		
PS	Disclosure; Col 49-56; 45pp; English.	
XX		
CC	The invention relates to an isolated DNA comprising the genome of an	
CC	enterically transmitted nonA/nonB (ET-NANB) viral hepatitis agent (also	
CC	referred as HEV). The DNA sequences or their fragments are useful in	
CC	preparing ET-NANB viral proteins and as probes for virus detection. These	
CC	are particularly useful in diagnosing infection by an enterically	
CC	transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.	
CC	The present sequence represents a ET-NANB viral DNA sequence from HEV-	
CC	Burma strain	
XX		
SQ	Sequence 7195 BP; 1221 A; 2312 C; 1868 G; 1794 T; 0 U; 0 Other;	
	Query Match	100.0%; Score 1983; DB 4; Length 7195;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 1983; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	ATGGCGCCTCGGCTATTTTGTGCTCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
Db	5147	ATGGCGCCTCGGCTATTTTGTGCTCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA 5206
QY	61	CCGCGCGGTGACCGGTGATTTCTAGCCCTTCGCAATCCCTATATTCACCAACCAAC 180
Db	5207	CCGCGCGGTGACCGGTGATTTCTAGCCCTTCGCAATCCCTATATTCACCAACCAAC 5326
QY	121	TTCTGGGGTGACCGGTGATTTCTAGCCCTTCGCAATCCCTATATTCACCAACCAAC 180
Db	5267	TTCTGGGGTGACCGGTGATTTCTAGCCCTTCGCAATCCCTATATTCACCAACCAAC 5326
QY	181	CCCTTCGCCCCCGATGTCACCGGTGCGGCGGGCTGGACCTCGTGTTCGCCAACCGCC 240
Db	5327	CCCTTCGCCCCCGATGTCACCGGTGCGGCGGGCTGGACCTCGTGTTCGCCAACCGCC 5386
QY	241	CGACCACTCGGCTCCGCTTGGCGGTGACCGGCCAGCGCCCGCGCGGTGCTCAGCGTCGT 300

Db 5387 CGACCACTCGGCTCCGCTTGGGTGACAGGCGCCAGCGCCGCTGCGCTCACGTCT 5446
QY 301 AGACCTTACCACAGCTGGGCGCGCGCTAAACCGCGGTGCTCCGCGCCATGACACCCCG 360
Db 5447 AGACCTTACCACAGCTGGGCGCGCGCTAAACCGCGGTGCTCCGCGCCATGACACCCCG 5506
QY 361 CAGTGCCTGATGTCAGCTCCCGCGGCGCACTTTGGCGGCGAGTAACTATCAACA 420
Db 5507 CCAGTGCCTGATGTCAGCTCCCGCGGCGCACTTTGGCGGCGAGTAACTATCAACA 5566
QY 421 TCTCCCTTACCTCTTCCGCGGCGCACTTAACTGCTTCTTATGCGCGCCCTCTT 480
Db 5567 TCTCCCTTACCTCTTCCGCGGCGCACTTAACTGCTTCTTATGCGCGCCCTCTT 5626
QY 481 AGTCCGCTTTTACCCCTTACAGCAGCGCACTTAACTATGCGGCGAGGCTTCT 540
Db 5627 AGTCCGCTTTTACCCCTTACAGCAGCGCACTTAACTATGCGGCGAGGCTTCT 5686
QY 541 AATTATGCCAGTACCGGTTGCGGTCGACCAATCGTTACCGCGCTGCTCCCAAT 600
Db 5687 AATTATGCCAGTACCGGTTGCGGTCGACCAATCGTTACCGCGCTGCTCCCAAT 5746
QY 601 GCTGTGCGGCTTACGCGCATCTCCATCTCATTTCTGGCCACAGACCACTCCCGAGC 660
Db 5747 GCTGTGCGGCTTACGCGCATCTCCATCTCATTTCTGGCCACAGACCACTCCCGAGC 5806
QY 661 TCCGTTGATATGAATCAATAAECTCGACGATGTTGATATTTAGTCCAGCGCGGATA 720
Db 5807 TCCGTTGATATGAATCAATAAECTCGACGATGTTGATATTTAGTCCAGCGCGGATA 5866
QY 721 GCCTGTGAGCTTGTGATCCCAAGTGAGGCTACACTATCGTAACAGGCTGCGCTCC 780
Db 5867 GCCTGTGAGCTTGTGATCCCAAGTGAGGCTACACTATCGTAACAGGCTGCGCTCC 5926
QY 781 GTCGAGACCTCTGGGCTGGCTGAGGAGGCTACCTCTGCTCTTGTATGCTTTGCATA 840
Db 5927 GTCGAGACCTCTGGGCTGGCTGAGGAGGCTACCTCTGCTCTTGTATGCTTTGCATA 5986
QY 841 CATGGCTCACTCGTAAATTCCTATATTAACCTTATACCGCTATACCGGTGCGCTGTTG 900
Db 5987 CATGGCTCACTCGTAAATTCCTATATTAACCTTATACCGGTGCGCTGTTG 6046
QY 901 GACTTTCGCTTGTGATGCTTTCGCAACCTTACCCCGGTACACCAATACCGGCTC 960
Db 6047 GACTTTCGCTTGTGATGCTTTCGCAACCTTACCCCGGTACACCAATACCGGCTC 6106
QY 961 TCCGCTTATTCAGCACTGCTCGCCAGCGCTTCTGTCGCGTGGGAGCGGACTGCCGAG 1020
Db 6107 TCCGCTTATTCAGCACTGCTCGCCAGCGCTTCTGTCGCGTGGGAGCGGACTGCCGAG 6166
QY 1021 CTCACCAAGCGCTGCTACCGCTTATGAGGAGCTCTATTTTACTAGTACTAATGGT 1080
Db 6167 CTCACCAAGCGCTGCTACCGCTTATGAGGAGCTCTATTTTACTAGTACTAATGGT 6226
QY 1081 GTCGCTGAGATGCGCGGGATAGCCTTCACTGTTTCACTTGTGACACTCTGCTT 1140
Db 6227 GTCGCTGAGATGCGCGGGATAGCCTTCACTGTTTCACTTGTGACACTCTGCTT 6286
QY 1141 GCGCGCTTCCGACAGAAATGATTTGCTGCGCTGGTGCCAGCTGTTCTACTCCCGTCCC 1200
Db 6287 GCGCGCTTCCGACAGAAATGATTTGCTGCGCTGGTGCCAGCTGTTCTACTCCCGTCCC 6346
QY 1201 GTTGTCTCAGCAATGGCGAGCGGCTTAAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 6347 GTTGTCTCAGCAATGGCGAGCGGCTTAAAGTTGTATACATCTGTAGAGAATGCTCAG 6406
QY 1261 CAGGATAGGTTATGCAATCCCGCATGACATGACCTCGGAGAACTCTGCTGCTGTTAT 1320
Db 6407 CAGGATAGGTTATGCAATCCCGCATGACATGACCTCGGAGAACTCTGCTGCTGTTAT 6466
QY 1321 CAGGATATGATAACCAATGAAACAAGATGGCGGAGCGCTTCTCGAGCCCATCGCGC 1380

Db 6467 CAGGATATGATAACCAATGAAACAAGATGCGCGCTTCTCAGCCCCATCGCGC 6526
QY 1381 CCTTCTCTGCTCTCGAGCTAATGATGCTTTGCTCTCTCTCACCGCTGCGGAGTAT 1440
Db 6527 CCTTCTCTGCTCTCGAGCTAATGATGCTTTGCTCTCTCTCACCGCTGCGGAGTAT 6586
QY 1441 GACCACTCCACTTATGGCTCTTTCGACTGGCCCAAGTTTATGTTCTGACTCTGAGCCTTG 1500
Db 6587 GACCACTCCACTTATGGCTCTTTCGACTGGCCCAAGTTTATGTTCTGACTCTGAGCCTTG 6646
QY 1501 GTTATGTTGCGACCGCGCGAGCGCTTGGCGGCTCGATTGGACCAAGGTCACA 1560
Db 6647 GTTATGTTGCGACCGCGCGAGCGCTTGGCGGCTCGATTGGACCAAGGTCACA 6706
QY 1561 CTTGACGGTGGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGC 1620
Db 6707 CTTGACGGTGGCGCCCTCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGC 6766
QY 1621 CTTCCGCGTAAAGCTCTTCTTCTGGGAGGCGAGGCAAACTAAAGCCGGGTACCTTATAAT 1680
Db 6767 CTTCCGCGTAAAGCTCTTCTTCTGGGAGGCGAGGCAAACTAAAGCCGGGTACCTTATAAT 6826
QY 1681 TATAACCACTCTAGCGACCAACTGCTTGTGAGAAATGCGCGCGGCACTCGGCTCGCT 1740
Db 6827 TATAACCACTCTAGCGACCAACTGCTTGTGAGAAATGCGCGCGGCACTCGGCTCGCT 6886
QY 1741 ATTTCCACTTACACCACTAGCCTGGGTGCTGGTCTCTCCATTTCTGCGGTTGCGGTT 1800
Db 6887 ATTTCCACTTACACCACTAGCCTGGGTGCTGGTCTCTCCATTTCTGCGGTTGCGGTT 6946
QY 1801 TTAGCCCCCACTCTCGGCTAGCAATTTGAGGATACCTTGGACTACCTGCCCCGCC 1860
Db 6947 TTAGCCCCCACTCTCGGCTAGCAATTTGAGGATACCTTGGACTACCTGCCCCGCC 7006
QY 1861 CATACTTTGATGATTTCTGCCAGAGTGGCGCCCTTCAAGGCTGCGCTTTC 1920
Db 7007 CATACTTTGATGATTTCTGCCAGAGTGGCGCCCTTCAAGGCTGCGCTTTC 7066
QY 1921 CAGTCTACTCTGCTGAGCTTTCAGCGCTTAAAGTGAAGTGGGTAATACTCGGAGTTG 1980
Db 7067 CAGTCTACTCTGCTGAGCTTTCAGCGCTTAAAGTGAAGTGGGTAATACTCGGAGTTG 7126
QY 1981 TAG 1983
Db 7127 TAG 7129

RESULT 9
AAL50386
ID AAL50386 standard; DNA; 7195 BP.
XX
AC AAL50386;
XX
DT 08-NOV-2002 (first entry)
XX
DE Hepatitis E virus (Burma strain) gene sequence.
XX
KW HEV; enterically transmitted nonA/nonB hepatitis viral agent;
Burma strain; bile; ds; gene; ORF1; ORF2; ORF3.
XX
OS Hepatitis E virus.
XX
FH Key Location/Qualifiers
CDS 28..5109
FT /*tag= a
FT /product= "Hepatitis E virus (Burma strain) ORF1 protein"
CDS 5106..5477
FT /*tag= c
FT /product= "Hepatitis E virus (Burma strain) ORF3 protein"
CDS 5147..7129
FT /*tag= b
FT /product= "Hepatitis E virus (Burma strain) ORF2 protein"
XX

```
PN US6379891-B1.
XX
PD 30-APR-2002.
XX
PF 19-APR-2000; 2000US-00553427.
XX
PR 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-00336672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 15-APR-1991; 91US-00681078.
PR 25-JUL-1994; 94US-00279823.
PR 07-JUN-1995; 95US-00478507.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam A, Fry KE;
XX
XX WPI; 2002-517277/55.
XX P-PSDB; AA015698, AA015699, AA015700.
XX
XX Detecting the presence of enterically transmitted nonA/nonB hepatitis
PT viral (HEV) agents in bile samples from infected humans and monkeys using
PT polymerase chain reaction.
XX
XX Disclosure; Col 13-24; 61pp; English.
PS
XX
XX The invention comprises a method for detecting the presence of
CC enterically transmitted nonA/nonB hepatitis viral (HEV) agents in a
CC sample and isolating HEV agents or nucleic acid fragments produced by the
CC agent. The method utilizes PCR - using bile from a human or cynomolgus
CC monkey actively infected with HEV as a source of the agent. The method of
CC the invention is used for detecting the presence of a viral agent in a
CC sample of cultured cells infected with the agent and isolating
CC enterically transmitted nonA/nonB HEV agents or nucleic acid fragments
CC produced by the agent. The present DNA sequence was isolated from a
CC Burmese strain of the Hepatitis E virus, the present DNA sequence has
CC three open reading frames ORF1, ORF2 and ORF3
XX
XX Sequence 7195 BP; 1221 A; 2312 C; 1868 G; 1794 T; 0 U; 0 Other;
SQ
Query Match
Best Local Similarity 100.0%; Score 1983; DB 6; Length 7195;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCCCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCGCGCCA 5206
DB 5147 ATGCGCCCTCGGCCCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCGCGCCA 5206
QY 61 CCGCCCGGTCAGCGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCTTCCGCGGTGT 120
DB 5207 CCGCCCGGTCAGCGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCTTCCGCGGTGT 5266
QY 121 TTCTGGGGTGACGGGGTGATTTCTCAGCCCTTCGCAATCCCTATATTTCATCCAAACCAAC 180
DB 5267 TTCTGGGGTGACGGGGTGATTTCTCAGCCCTTCGCAATCCCTATATTTCATCCAAACCAAC 5326
QY 181 CCCTTCGCCCCCGATGTACCGCTCGCGCGCGGGGTGACCTCTGTTTCGCAACCCCGCC 240
DB 5327 CCCTTCGCCCCCGATGTACCGCTCGCGCGCGGGGTGACCTCTGTTTCGCAACCCCGCC 5386
QY 241 CGAACCTACGAGCTGGGGCGCGCTGACAGGCCCGCGCGCGCGCTTCCGCTCAGTGT 300
DB 5387 CGAACCTACGAGCTGGGGCGCGCTGACAGGCCCGCGCGCGCGCTTCCGCTCAGTGT 5446
QY 301 AGACCTACCAAGCTGGGGCGCGCTAACCGGGCTCGCTCCGCGCGCGCTGACACCCCG 360
DB 5447 AGACCTACCAAGCTGGGGCGCGCTAACCGGGCTCGCTCCGCGCGCGCTGACACCCCG 5506
QY 361 CCAGTGCCTGTATGTGACTCCGCGCGGCCATCTTTCGCGCGCGAGTATTAACCTATCAACA 420
DB
```


CC cynomolgus monkeys infected with the Burma strain of ET-NANB. Both
 CC strands of Etl.1 were sequenced. Identity of the sequence with sequences
 CC in etiologic agents has been confirmed by locating a similar sequence in
 CC a viral strain isolated in Burma. The sequence of the Burma strain is
 CC given here. See AA014410 for Etl.1. (Updated on 25-MAR-2003 to correct PA
 CC field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-
 CC 2003 to correct OS field.)
 XX

SQ Sequence 7195 BP; 1221 A; 2309 C; 1859 G; 1806 T; 0 U; 0 Other;

Query Match 99.4%; Score 1970.2; DB 2; Length 7195;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1975; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy	1	ATGGGCGCTCGGCTATTTGTTGCTGCTCTCATGTTTGTGCTATGCTATGCTGCGCGCCA	60
Db	5147	ATGGGCGCTCGGCTATTTGTTGCTGCTCTCATGTTTGTGCTATGCTATGCTGCGCGCCA	5206
Qy	61	CCGCGCGCTCAGCGCTCGCGCGCGCTGCTGGGCGCGCAGCGCGGTTCCGGCGGTGGT	120
Db	5207	CCGCGCGCTCAGCGCTCGCGCGCGCTGCTGGGCGCGCAGCGCGGTTCCGGCGGTGGT	5266
Qy	121	TTCTGGGGTCAACCGGTTGATCTCAGCCCTTCGAATCCCTATATTCATCCAAACCAAC	180
Db	5267	TTCTGGGGTCAACCGGTTGATCTCAGCCCTTCGAATCCCTATATTCATCCAAACCAAC	5326
Qy	181	CCCTTGGCGCGGATGTCAGCGTTCGCGCGCGGCTGGAGCTCGTGTTCGCCACCGCC	240
Db	5327	CCCTTGGCGCGGATGTCAGCGTTCGCGCGCGGCTGGAGCTCGTGTTCGCCACCGCTCC	5386
Qy	241	CGACCACTCGGCTCCGCTTGGCGTGACAGCGCCAGCGCCCGCGCTTGCTCAGCTCGT	300
Db	5387	CGACCACTCGGCTCCGCTTGGCGTGACAGCGCCAGCGCCCGCGCTTGCTCAGCTCGT	5446
Qy	301	AGACCTACCAAGCTGGGGCGCGCGCTAAACCGCGTGGCTTCGCGGCCATGACACCGCG	360
Db	5447	AGACCTACCAAGCTGGGGCGCGCGCTAAACCGCGTGGCTTCGCGGCCATGACACCGCG	5506
Qy	361	CCAGTGCCTGATGTCGACTCCGCGCGCGCATCTTGGCGCGCAGTATACCTATCAACA	420
Db	5507	CCAGTGCCTGATGTCGACTCCGCGCGCGCATCTTGGCGCGCAGTATACCTATCAACA	5566
Qy	421	TCTCCCTTACCTCTCCGCTGGCCACCGGCACTAACCTGGTTCCTTATGCGCCCTCTT	480
Db	5567	TCTCCCTTACCTCTCCGCTGGCCACCGGCACTAACCTGGTTCCTTATGCGCCCTCTT	5626
Qy	481	AGTCCGCTTTTACCTTTCAGGACGGGACCAATACCAATATATATGCGGACGGAAGTCT	540
Db	5627	AGTCCGCTTTTACCTTTCAGGACGGGACCAATACCAATATATATGCGGACGGAAGTCT	5686
Qy	541	AATTTATGCCAGTACCGGGTTGCCGTGCCAATCCGTTACCGCGCTGGTCCCAAT	600
Db	5687	AATTTATGCCAGTACCGGGTTGCCGTGCCAATCCGTTACCGCGCTGGTCCCAAT	5746
Qy	601	GCTGTCCGCGGTATACCGCATCTCCATCTCATTTCTGGCCACAGACCAACCCCGGACG	660
Db	5747	GCTGTCCGCGGTATACCGCATCTCCATCTCATTTCTGGCCACAGACCAACCCCGGACG	5806
Qy	661	TCGCTGATATGAATTCATTAACCTCGACGATGTTGATATTTAGTCCAGCCCGGCATA	720
Db	5807	TCGCTGATATGAATTCATTAACCTCGACGATGTTGATATTTAGTCCAGCCCGGCATA	5866
Qy	721	GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGTCC	780
Db	5867	GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGTCC	5926
Qy	781	GTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCTGGTCTTGTATGCTTTGCATA	840
Db	5927	GTCGAGACCTCTGTGGTGGCTGAGGAGGAGCTACCTCTGGTCTTGTATGCTTTGCATA	5986
Qy	841	CATGGCTCACTGTAATTCCTATATCACTAATACCCCTATACGGTGGCTCGGCTGTTG	900
Db	5987	CATGGCTCACTGTAATTCCTATATCACTAATACCCCTATACGGTGGCTCGGCTGTTG	6046

Qy	901	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGCTC	960
Db	6047	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGCTC	6106
Qy	961	TCCCGTTATTCAGCACTGCTCGCCACCGCTTTCGTCGCGGTGGAGCGGACTGCGGAG	1020
Db	6107	TCCCGTTATTCAGCACTGCTCGCCACCGCTTTCGTCGCGGTGGAGCGGACTGCGGAG	6166
Qy	1021	CTCACCACCAAGCTGCTACCCGCTTATGAAGAGCTCTATTTACTAGTACTAATGGT	1080
Db	6167	CTCACCACCAAGCTGCTACCCGCTTATGAAGAGCTCTATTTACTAGTACTAATGGT	6226
Qy	1081	GTCCGTGAGATCGCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTT	1140
Db	6227	GTCCGTGAGATCGCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTT	6286
Qy	1141	GGCGGCTCGCGCAGAAATTTGATTCGTCGCGTGGTGGCCAGCTGTTCTACTCCGTCCT	1200
Db	6287	GGCGTCTCGCGCAGAAATTTGATTCGTCGCGTGGTGGCCAGCTTCTACTCCGTCCT	6346
Qy	1201	GTGTCTCAGCAATGCGAGCGGCTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG	1260
Db	6347	GTGTCTCAGCAATGCGAGCGGCTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG	6406
Qy	1261	CAGGATAAGGGTATTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGGTTATT	1320
Db	6407	CAGGATAAGGGTATTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGGTTATT	6466
Qy	1321	CAGGATATGATTAACCAATGAAGTTCGCGCGAGCGCTTCTCAGGCCCATCGCGC	1380
Db	6467	CAGGATATGATTAACCAATGAAGTTCGCGCGAGCGCTTCTCAGGCCCATCGCGC	6526
Qy	1381	CCCTTCTCTGCTTCGAGCTTAATGATGCTTTGGCTCTCTCTCAGCGTGGCGAGTAT	1440
Db	6527	CCCTTCTCTGCTTCGAGCTTAATGATGCTTTGGCTCTCTCTCAGCGTGGCGAGTAT	6586
Qy	1441	GACCACTTCACTTATGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG	1500
Db	6587	GACCACTTCACTTATGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG	6646
Qy	1501	GTTAATGTTGGGACCGCGCGCGCGCTTTCGCGGTGGCTCGTTCGATTCGACCAAGGTACA	1560
Db	6647	GTTAATGTTGGGACCGCGCGCGCGCTTTCGCGGTGGCTCGTTCGATTCGACCAAGGTACA	6706
Qy	1561	CTTGACGCTCGCCCGCTCTCCACCATCCAGCACTCGAAGACCTTCTTGTCTGCTGCG	1620
Db	6707	CTTGACGCTCGCCCGCTCTCCACCATCCAGCACTCGAAGACCTTCTTGTCTGCTGCG	6766
Qy	1621	CTCCGCGTTAAGCTCTCTTTCTGGGAGGCGAGGCAACAATAAGCCGGGTACCTTTAAT	1680
Db	6767	CTCCGCGTTAAGCTCTCTTTCTGGGAGGCGAGGCAACAATAAGCCGGGTACCTTTAAT	6826
Qy	1681	TATAACACCACTGCTAGGACCAATCTGTTGAGAAATGCGCGCGGCGGCGGCTGCT	1740
Db	6827	TATAACACCACTGCTAGGACCAATCTGTTGAGAAATGCGCGCGGCGGCGGCTGCT	6886
Qy	1741	ATTTTCCACTTACACCACTAGCTGGGTGCTGCTCCCTCTCCATTTCTGCGGTGGCGCTT	1800
Db	6887	ATTTTCCACTTACACCACTAGCTGGGTGCTGCTCCCTCTCCATTTCTGCGGTGGCGCTT	6946
Qy	1801	TTAGCCCCCACTCTGCGCTAGCAATGCTTGAAGATACCTTGAAGTACCTCCCTCCCGGCC	1860
Db	6947	TTAGCCCCCACTCTGCGCTAGCAATGCTTGAAGTACCTTGAAGTACCTCCCTCCCGGCC	7006
Qy	1861	CATACTTTTATGATTTCTGCCAGAGTGGCGCCCTTGGCTTCCAGGCTGGCTTTC	1920
Db	7007	CATACTTTTATGATTTCTGCCAGAGTGGCGCCCTTGGCTTCCAGGCTGGCTTTC	7066
Qy	1921	CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTGGAATCTCGGAGTTG	1980
Db	7067	CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTGGAATCTCGGAGTTG	7126

QY	1981 TAG 1983	301	AGACCTACCAAGCTGGGGCGCGCGCTAAACCGCGTGGCTCGCTCCGCGCCATGACACCCCG	360
DB	7127 TAG 7129	DB		
RESULT 12				
AAV61688		301	AGACCTACCAAGCTGGGGCGCGCGCTAAACCGCGTGGCTCGCTCCGCGCCATGACACCCCG	360
ID	AAV61688 standard; DNA; 1983 BP.	361	CCAGTGCCCTGATGTCGACTCTCGCGGCGCATCTTGCGCGCGGAGTATAATCTATCAACA	420
XX	AAV61688;	DB		
AC		361	CCAGTGCCCTGATGTCGACTCTCGCGGCGCATCTTGCGCGCGGAGTATAATCTATCAACA	420
XX		QY		
DT	03-DEC-1998 (first entry)	421	TCTCCCTTACCTCTTCGCTGGCGCACCGGACCTAACCTGGTCTTTTATGCGCGCCCTCTT	480
XX		DB		
DE	Hepatitis E virus hollow particle protein DNA #2.	421	TCTCCCTTACCTCTTCGCTGGCGCACCGGACCTAACCTGGTCTTTTATGCGCGCCCTCTT	480
XX		QY		
KW	Hollow particle protein; virus; antibody; detection; immunoassay;	481	AGTCGCGCTTTTACCCCTTCAGGACGCGACCAATACCATATAATGCGCCACGGAAGCTTCT	540
XX	infection; ss.	DB		
XX	Hepatitis virus.	481	AGTCGCGCTTTTACCCCTTCAGGACGCGACCAATACCATATAATGCGCCACGGAAGCTTCT	540
OS		QY		
XX		541	AATTATGCCAGTACCGGTTGCCGTCGACCAATCCGTTACCGCGCGTGGTCCCAAT	600
PN	JP10234383-A.	DB		
PD	08-SEP-1998.	541	AATTATGCCAGTACCGGTTGCCGTCGACCAATCCGTTACCGCGCGTGGTCCCAAT	600
XX		QY		
PF	28-FEB-1997; 97JP-00062445.	601	GCTGTCGGGCTTACGCCATCTCCATCTCATCTCTGCGCACAGACCAACACCCCGGAG	660
XX		DB		
PR	28-FEB-1997; 97JP-00062445.	601	GCTGTCGGGCTTACGCCATCTCCATCTCATCTCTGCGCACAGACCAACACCCCGGAG	660
XX	(DENK-) DENKA SEIKEN KK.	661	TCCGTTGATGATTAATTAACCTCGACGATGTTGTTATTTAGTCCAGGCCCGGCATA	720
PA	(KOKO-) KOKURITSU YOHIO EISEI KENKYUSHO.	DB		
XX	WPI; 1998-535037/46.	661	TCCGTTGATGATTAATTAACCTCGACGATGTTGTTATTTAGTCCAGGCCCGGCATA	720
DR	P-PSDB; AAW76369.	QY		
XX		721	GCTCTGAGCTTGTGATGCCAAGTGGCGCTTACACTATCGTAAACGAGCTGGCGCTCC	780
PT	Hepatitis E virus hollow particle poly:peptide(s) and nucleic acids	DB		
PT	encoding it - useful for more accurate detection of HEV in samples, using	721	GCTCTGAGCTTGTGATGCCAAGTGGCGCTTACACTATCGTAAACGAGCTGGCGCTCC	780
PT	immuno-assays and nucleic acid hybridisation.	QY		
XX	Claim 13; Page 24-26; 29pp; Japanese.	781	GTCGAGACTCTGGGCTGGCTGAGGAGGAGGCTACTCTGCTGTTGTTATGCTTGCATA	840
XX		DB		
CC	This sequence encodes a Hepatitis E viral hollow particle protein. This	781	GTCGAGACTCTGGGCTGGCTGAGGAGGAGGCTACTCTGCTGTTGTTATGCTTGCATA	840
CC	polypeptide can be used to raise antibodies to detect HEV infection in	QY		
CC	samples, e.g. by immuno-assay based techniques, and the nucleic acid can	DB		
CC	be used for the same in nucleic acid hybridisation assays. The	841	CATGGCTCACTCGTAAATTCCTATATAACACCTATACCGGTGCCCTCCGGGTGTTG	900
CC	polypeptides and nucleic acids allow more accurate detection of HEV than	DB		
CC	previously possible	841	CATGGCTCACTCGTAAATTCCTATATAACACCTATACCGGTGCCCTCCGGGTGTTG	900
XX		QY		
PS	Sequence 1983 BP; 342 A; 665 C; 485 G; 491 T; 0 U; 0 Other;	901	GACTTGGCTTGGCTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC	960
XX		DB		
CC	Query Match 98.3%; Score 1949.4; DB 2; Length 1983;	901	GACTTGGCTTGGCTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC	960
CC	Best Local Similarity 98.9%; Pred. No. 0;	QY		
CC	Matches 1962; Conservative 0; Mismatches 21; Indels 0; Gaps 0;	DB		
QY	1 ATGCGCCCTCGCCCTATTTTGTGCTGCTCCTCATGTTTTCGCTATGTCGCGGCCA	1081	GTCGCTGAGATCGCGCGGGGATAGCCCTCACCTGTTTCAACCTTGCTGACACTCTGCTT	1140
DB	60	DB		
QY	1 ATGCGCCCTCGCCCTATTTTGTGCTGCTCCTCATGTTTTCGCTATGTTGCGGCCA	1081	GTCGCTGAGATCGCGCGGGGATAGCCCTCACCTGTTTAACTTGTGCTGACACTCTGCTT	1140
DB	60	QY		
QY	61 CGCGCCGCTCAGCGCTGCGCGCGCTGCTGGCGCGCGCGAGCGGCTTCCGCGGTGT	1141	GSCGCGCTTCCGACAGAAATGATTTTCGCTGGCTGGTGGCGAGCTGTTCTACTCCGTC	1200
DB	120	DB		
QY	61 CGCGCCGCTCAGCGCTGCGCGCGCTGCTGGCGCGCGAGCGGCTTCCGCGGTGT	1141	GSCGCGCTTCCGACAGAAATGATTTTCGCTGGCTGGTGGCGAGCTGTTCTACTCCGTC	1200
DB	120	QY		
QY	121 TTCGGGGTGACCGGGTGTGATTTCTCAGCCCTTTCGCAATCCCTATATTCACCAAC	1201	GTGTCTCAGCCCAATGCGGAGCGGAGCTGTTAAGTTGTATACATCTGTAGAGAAATGCT	1260
DB	180	DB		
QY	121 TTCGGGGTGACCGGGTGTGATTTCTCAGCCCTTTCGCAATCCCTATATTCACCA	1201	GTGTCTCAGCCCAATGCGGAGCGGAGCTGTTAAGTTGTATACATCTGTAGAGAAATGCT	1260
DB	180	QY		
QY	181 CCCTTCGCCCCGATGTACACCGTGGCGCGGGGTGACCTCGTGTTCGCAACCCGCG	1261	CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT	1320
DB	240	DB		
QY	181 CCCTTCGCCCCGATGTACACCGTGGCGCGGGGTGACCTCGTGTTCGCAACCCGCG	1261	CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT	1320
DB	240	QY		
QY	241 CGACCACTCGGCTCCGCTTGGCTGACAGGCGCGCGCGGCTTCCGAGCCCATCGCG	1321	CAGGATTATGATAACCAACATGAAAGATCGGCGAGCGCTTCTTCAGGCCCATCGCGC	1380
DB	300	DB		
QY	241 CGACCACTCGGCTCCGCTTGGCTGACAGGCGCGCGCGGCTTCCGAGCCCATCGCG	1321	CAGGATTATGATAACCAACATGAAAGATCGGCGAGCGCTTCTTCAGGCCCATCGCGC	1380
DB	300	QY		
QY		1381	CCTTCTCTGCTCCTTCGAGCTAATGATGTGCTTTTGGCTCTCTCTCACCGCTGCCAGTAT	1440

Db 1381 CTTTTCTGTGCTTCGAGCTAATGATGCTTTGGCTCTCTCTACCGTGGCGAGTAT 1440
QY 1441 GACCACTCCACTTATGGCTCTTCGACTGGCCCAAGTTTATGTTTCTGACTCTGTGACTTGG 1500
Db 1441 GACCACTCCACTTATGGCTCTTCGACTGGCCCAAGTTTATGTTTCTGACTCTGTGACTTGG 1500
QY 1501 GTTAATGTTGCGACCGCGGCGAGCGGTTGCCGGTTCGATTCGATTCGACCAAGTTCACA 1560
Db 1501 GTTAATGTTGCGACCGCGGCGAGCGGTTGCCGGTTCGATTCGATTCGACCAAGTTCACA 1560
QY 1561 CTTGACGCTGCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTGCTCTGCGG 1620
Db 1561 CTTGACGCTGCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTTCTTTGCTCTGCGG 1620
QY 1621 CTCGCGGTAAGCTCTCTTTCTGGAGGCGAGGCAACAATAAGCCGGGTACCTTATAAT 1680
Db 1621 CTCGCGGTAAGCTCTCTTTCTGGAGGCGAGGCAACAATAAGCCGGGTACCTTATAAT 1680
QY 1681 TATAACACCACTGCTAGCGAACAATGCTTGTGAGAAATGCGCGGCGACCGGTCGCT 1740
Db 1681 TATAACACCACTGCTAGCGAACAATGCTTGTGAGAAATGCGCGGCGACCGGTCGCT 1740
QY 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTTGCCGTT 1800
Db 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTTGCCGTT 1800
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTGCTTGGAGTACCTTGGACTACCTTGGACTACCTTGG 1860
Db 1801 TTAGCCCCCACTCTGCGCTAGCATTGCTTGGAGTACCTTGGACTACCTTGGACTACCTTGG 1860
QY 1861 CATACTTTGATGATTTCTGCGCAGAGTGGCGCCCTTGGCTTCCAGGCTGCGCTTC 1920
Db 1861 CATACTTTGATGATTTCTGCGCAGAGTGGCGCCCTTGGCTTCCAGGCTGCGCTTC 1920
QY 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAAATCTCGGAGTTG 1980
Db 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAAATCTCGGAGTTA 1980
QY 1981 TAG 1983
Db 1981 TAG 1983

RESULT 13
AAV61690
ID AAV61690 standard; DNA; 7194 BP.

XX AC AAV61690;

XX DT 03-DEC-1998 (first entry)

XX DE Hepatitis E virus hollow particle protein DNA #4.

XX KW Hollow particle protein; virus; antibody; detection; immunoassay;
infection; ss.

XX OS Hepatitis virus.

XX PN JP10234383-A.

XX PD 08-SEP-1998.

XX PF 28-FEB-1997; 97JP-00062445.

XX PR 28-FEB-1997; 97JP-00062445.

XX PA (DENK-) DENKA SEIKEN KK.

XX PA (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.

XX DR WPI; 1998-535037/46.

XX PT Hepatitis E virus hollow particle poly:peptide(s) and nucleic acids

PT encoding it - useful for more accurate detection of HEV in samples, using
PT immuno-assays and nucleic acid hybridisation.
XX Disclosure; Page 27-29; 29pp; Japanese.
XX This sequence encodes a Hepatitis E viral hollow particle protein. This
CC polypeptide can be used to raise antibodies to detect HEV infection in
CC samples, e.g. by immuno-assay based techniques, and the nucleic acid can
CC be used for the same in nucleic acid hybridisation assays. The
CC polypeptides and nucleic acids allow more accurate detection of HEV than
CC previously possible
XX Sequence 7194 BP; 1217 A; 2312 C; 1871 G; 1794 T; 0 U; 0 Other;

Query Match 98.3%; Score 1949.4; DB 2; Length 7194;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1962; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 ATCGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
Db 5147 ATCGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTGTGCTATGTTGCGCGGCCA 5206
QY 61 CCGCCCGGTGACGCGTCTGCGCCCGTCTGCGCGCGGCGAGCGGCTTCCGGCGGTGGT 120
Db 5207 CCGCCCGGTGACGCGTCTGCGCCCGTCTGCGCGCGGCGAGCGGCTTCCGGCGGTGGT 5266
QY 121 TTCTGGGGTGACCGGCTTGAATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 5267 TTCTGGGGTGACCGGCTTGAATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 5326
QY 181 CCCTTGGCCCCGATGTACCGCTGGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 240
Db 5327 CCCTTGGCCCCGATGTACCGCTGGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 5386
QY 241 CGACCACTCTCGGCTCCGCTGGCTGACCGGCGCGGCGGCGGCTTGTGCTCAGCTCGT 300
Db 5387 CGACCACTCTCGGCTCCGCTGGCTGACCGGCGCGGCGGCGGCTTGTGCTCAGCTCGT 5446
QY 301 AGACCTACCACTGCGGCGCGGCGGCTTAACCGGCTCGCTCCGCGCGGCGGCGGCGG 360
Db 5447 AGACCTACCACTGCGGCGCGGCGGCTTAACCGGCTCGCTCCGCGCGGCGGCGGCGG 5506
QY 361 CCAGTGCTGATGTGCACTCCCGCGGCGGCGGCTTTCGCGCGGCGGCGGCTTATCAACA 420
Db 5507 CCAGTGCTGATGTGCACTCCCGCGGCGGCGGCTTTCGCGCGGCGGCGGCTTATCAACA 5566
QY 421 TCTCCGCTTACCTCTTCCGCTGGCGGCGGCGGCGGCTTCTTATGCGGCGGCTTCT 480
Db 5567 TCTCCGCTTACCTCTTCCGCTGGCGGCGGCGGCGGCTTCTTATGCGGCGGCTTCT 5626
QY 481 AGTCCGCTTTTACCCCTTTCAGGACGCGGCGGCGGCTTAACCGGCTCGCTCCGCGGCGGCGG 540
Db 5627 AGTCCGCTTTCACCTTTCAGGACGCGGCGGCGGCTTAACCGGCTCGCTCCGCGGCGGCGG 5686
QY 541 AATTATGCCAGTACCGGCTTTCGCGGCGGCGGCGGCTTACCGGCGGCTTTCGCGGCGGCGG 600
Db 5687 AATTATGCCAGTACCGGCTTTCGCGGCGGCGGCGGCTTACCGGCGGCTTTCGCGGCGGCGG 5746
QY 601 GCTGTGCGGCTTACCGGCTTTCGCGGCGGCGGCGGCTTACCGGCGGCTTTCGCGGCGGCGG 660
Db 5747 GCTGTGCGGCTTACCGGCTTTCGCGGCGGCGGCGGCTTACCGGCGGCTTTCGCGGCGGCGG 5806
QY 661 TCGGCTGATGATGAATTAATCAATTAATCAATTAATCAATTAATCAATTAATCAATTAAT 720
Db 5807 TCGGCTGATGATGAATTAATCAATTAATCAATTAATCAATTAATCAATTAATCAATTAAT 5866
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGGCTTACCTATCGTAAACCAAGGCTGGCGCTCC 780
Db 5867 GCCTCTGAGCTTGTGATCCCAAGTGAGCGGCTTACCTATCGTAAACCAAGGCTGGCGCTCC 5926
QY 781 GTCGAGACCTCTGGGCTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Db 5927 GTCGAGACCTCTGGGCTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5986

841 CATGGCTCACTCGTAAATTCCTATATACTAATACACCCCTATACCGGTCCCTCGGGCTGTG 900
5987 CATGGCTCACTCGTAAATTCCTATATACTAATACACCCCTATACCGGTCCCTCGGGCTGTG 6046
901 GACTTTGCCCTTGAGCTTGAGTTTCCGAACTTTACCCCGGTAAACCAATACGCGGTC 960
6047 GACTTTGCCCTTGAGCTTGAGTTTCCGAACTTTACCCCGGTAAACCAATACGCGGTC 6106
961 TCCCGTTATTCAGACACTGCTCCGACCGCTTCGTGCGGTGCGGACGGGACTGCCGAG 1020
6107 TCCCGTTATTCAGACACTGCTCCGACCGCTTCGTGCGGTGCGGACGGGACTGCCGAG 6166
1021 CTCACCAACACCGCTGCTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGT 1080
6167 CTCACCAACACCGCTGCTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGT 6226
1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCCCTGTTCAACCTTCTGACACTCTGCTT 1140
6227 GTCGGTGAGATCGGCGCGGATAGCCCTCACCCCTGTTTAACTTCTGACACTCTGCTT 6286
1141 GCGGCGCTGCCACAGAAATGATTTGTCGGCTGGTGGCAGCTGTCTACTCCCGTCC 1200
6287 GCGGCGCTGCCACAGAAATGATTTGTCGGCTGGTGGCAGCTGTCTACTCCCGTCC 6346
1201 GTTGTCTCAGCAATGCGGAGCGACTGTTAAGTTGTATACATCTCTAGAGATGCTCAG 1260
6347 GTTGTCTCAGCAATGCGGAGCGACTGTTAAGTTGTATACATCTCTAGAGATGCTCAG 6406
1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTCGCGGAGATCTCGTGTGTTATT 1320
6407 CAGGATAAGGGTATTGCAATCCCGCATGACATTCGCGGAGATCTCGTGTGTTATT 6466
1321 CAGGATTATGATAACCAATGAAACAGATGCGCGACGCTTCTCCAGCCCGATCGCGC 1380
6467 CAGGATTATGATAACCAATGAAACAGATGCGCGACGCTTCTCCAGCCCGATCGCGC 6526
1381 CTTTCTCTGCTCTTCGAGCTAATGATGCTTTTGGCTCTCTCTCAGCGCTGCCAGTAT 1440
6527 CTTTCTCTGCTCTTCGAGCTAATGATGCTTTTGGCTCTCTCTCAGCGCTGCCAGTAT 6586
1441 GACCACTCACTATTGGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
6587 GACCACTCACTATTGGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG 6646
1501 GTTATGTTGACCGCGCGCAGCGCTGTCGCGTGCCTCGATGAGTGAACAGGTACA 1560
6647 GTTATGTTGACCGCGCGCAGCGCTGTCGCGTGCCTCGATGAGTGAACAGGTACA 6706
1561 CTTGACGGTTCGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTGCTCGCG 1620
6707 CTTGACGGTTCGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTGCTCGCG 6766
1621 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGCAGCAACAACTAAAGCCGGGTACCTTATAAT 1680
6767 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGCAGCAACAACTAAAGCCGGGTACCTTATAAT 6826
1681 TATAACCACTGCTAGCAGCACTGCTTTCGAGATGCGCGCGGACCGGTCGCT 1740
6827 TATAACCACTGCTAGCAGCACTGCTTTCGAGATGCGCGCGGACCGGTCGCT 6886
1741 ATTTCACCTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
6887 ATTTCACCTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6946
1801 TTAGCCCCCACTCTGCGCTAGCACTTGTGAGGATACCTTGGACTACCTCCCGCGCGCC 1860
6947 TTAGCCCCCACTCTGCGCTAGCACTTGTGAGGATACCTTGGACTACCTCCCGCGCGCC 7006
1861 CATACCTTTGATGATTTCTGCCAGAGTGCGCCCGCTTGGCTTGAAGGCTGCGCTTTC 1920
7007 CATACCTTTGATGATTTCTGCCAGAGTGCGCCCGCTTGGCTTGAAGGCTGCGCTTTC 7066

1921 CAGTCTACTGTCGCTGAGCTTCAGCGCTTAAAGATGAAGTGGTAAACTCGGAGTTG 1980
7067 CAGTCTACTGTCGCTGAGCTTCAGCGCTTAAAGATGAAGTGGTAAACTCGGAGTTA 7126
1981 TAG 1983
7127 TAG 7129
RESULT 14
AAQ86592
ID AAQ86592 standard; DNA; 1984 BP.
XX AAQ86592;
AC AAQ86592;
XX
DT 25-MAR-2003 (revised)
DT 28-SEP-1995 (first entry)
XX
DE Hepatitis E virus ORF2.
XX
XX HEV; ORF2; antigen; vaccine; immunogen; ss.
XX
OS Hepatitis E virus.
XX
FH Key Location/Qualifiers
FT CDS 1..1983
FT /*tag= a
XX
PN W09508632-A1.
XX
PD 30-MAR-1995.
XX
XX 23-SEP-1994; 94WO-AU0000572.
PF
XX 24-SEP-1993; 93AU-00001423.
PR 15-DEC-1993; 93AU-00002964.
XX
XX (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
XX Anderson DA, Locarnini SA, Torresi J, Li F, Hui Z;
PI
XX WPI; 1995-139601/18.
DR P-PSDB; AAR70323.
XX
PT Antigens of hepatitis E virus (HEV) - selectively immuno-reactive to
PT convalescent and/or acute phase circulating antibodies to HEV.
XX
PS Disclosure; Page 39-41; 78pp; English.
XX
CC RNA from an HEV strain isolated in the Xinjiang region of China was
CC subjected to RT-PCR to obtain fragments corresp. to ORF2 (given in
CC AAQ86592) and ORF3 (AAQ86593) that encoded antigenic proteins (AAR70323-
CC 24, respectively). DNA fragments were manipulated into pGEX vectors for
CC production of GST fusion proteins in E. coli. Applications include HEV
CC infection diagnosis, therapy and vaccine development. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
SQ Sequence 1984 BP; 347 A; 666 C; 481 G; 490 T; 0 U; 0 Other;
Query Match 91.0%; Score 1805.4; DB 2; Length 1984;
Best Local Similarity 94.4%; Pred No. 0;
Matches 1872; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTTTCCTATGTCGCCCGCGCA 60
DB 1 ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTTTTCCTATGTCGCCCGCGCA 60
QY 61 CGCCCGGTGAGCGCTCTGGCGCGCGCTGTCGGCGCGCAGCGCGGTTCGGCGGTG 120
DB 61 CGCCCGGTGAGCGCTCTGGCGCGCGCTGTCGGCGCGCAGCGCGGTTCGGCGGTG 120
QY 121 TTCTGGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTTCACCAACAC 180

Db	121	TTCTGGGGTGACCGGGTTGATTTCTAGCCCTTCGCAATCCCTATATTATTCATCAACCAAC	180
Qy	181	CCCTTCCGCCCGCATGTCAACGCTGGCGCGCGGGCTGGACCTCGTGTTCGCCAACCGGC	240
Db	181	CCCTTCCGCCCGCATGTCAACGCTGGCGCGCGGGCTGGACCTCGTGTTCGCCAACCGGC	240
Qy	241	CGAACCACTCGGCTCCGCTTGGGTGGACCAAGGCCAGCGCCCGCGCTTGCTCTACGTCGT	300
Db	241	CGAACCACTCGGCTCCGCTTGGGTGGACCAAGGCCAGCGCCCGCGCTTGCTCTACGTCGT	300
Qy	301	AGACCTACACAGCTGGGGCGGCGCGCTAACCGGGTGCCTCCGCGCCCATGACACCCCG	360
Db	301	AGACCTACCAAGCTGGGGCGGCGCGCTAACCGGGTGCCTCCGCGCCCATGACACCCCG	360
Qy	361	CCAGTGCTGATGTCGACTCCCGCGGGCCCATCTTTCGCGCGGAGTATAACCTATCAACA	420
Db	361	CCAGTGCTGATGTTGACTCCCGGGGGCCCATCTTTCGCGCGGAGTATAACCTATCAACA	420
Qy	421	TCTCCCTTACCTCTTCCGTGGGCACCGGCACTAAACCTGGTTCCTTTATGCGCGCCCTCTT	480
Db	421	TCTCCCTTACTTCTTCGTGGGCACCGGTACAACTTGGTTCCTATACGCGCTCCTCTT	480
Qy	481	AGTCGGCTTTTACCCCTTCAGGACGCGACCAATACCACTATATATGCGCAGGAAGCTCT	540
Db	481	AGCCCACTTTCACCCCTCCAGGACGCGACCAATACTATATATATGCGCAGGAAGCTCT	540
Qy	541	AAATTATGCCAGTACCGGTTGCCGTGGCCACCAATCCGTTACCGCGCGTGGTTCGCCAAT	600
Db	541	AAATTATGCCAGTACCGGTTGGTTCGTGGCCACCAATTCGCTACCGCGCGTGGTTCGCCA	600
Qy	601	GCTGTCCGCGGTTACGCCATCTCCATCTCAATTCGTGGCCACAGACCAACCCCGAGC	660
Db	601	GCTGTGTGTGCTACGCCATCTCCATCTCGTTCGTGGCCACAGACCAACCCCGAGC	660
Qy	661	TCCGTTGATATGAATTCATTAACCTCGACGGATGTTGATATTTAGTCCAGCCCGCAT	720
Db	661	TCCGTTGACATGAATTCATTAACCTCGACGGATGTTGATATTTAGTCCAGCCCGCAT	720
Qy	721	GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACAAAGCTCGCGCTCC	780
Db	721	GCCTCCGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACAAAGCTCGCGCTCT	780
Qy	781	GTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCTGCTCTGTTATGCTTTGCATA	840
Db	781	GTTTGAGACCTCCCGGGTGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	840
Qy	841	CATGGCTCACTCGTAATTCCTATCTAATAACACCTATACCGGTGCGCTCGGGCTGTTG	900
Db	841	CATGGCTCACCTGTAATTCCTATCTAATAACCTATACCGGTGCGCTCGGGCTGTTG	900
Qy	901	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGTC	960
Db	901	GACTTTGCCCTTCGAACTTGAGTTCGCAACCTTACCCCGGTAAACCAACGCGGTC	960
Qy	961	TCCCGTTATTCAGCACTGCTCGCCACCGCCTTCGTCCGCGTGGGACGCGGAGCTGCCG	1020
Db	961	TCCCGTTACTCCAGCACTGCTCGCCACCGCCTTCGTCCGCGTGGGACGCGGAGCTGCCG	1020
Qy	1021	CTCACCAACGCGGTGTACCCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGT	1080
Db	1021	CTTACCAACGCGGTGTACCCGCTTCATGAAGACCTCTATTTTACTAGTACTAATGGT	1080
Qy	1081	GTCCGTGAGATCGCGCGGGATAGCCCTCAACCTGTTCAACCTTGCTGACACTCTGCTT	1140
Db	1081	GTCCGTGAGATCGCGCGGTGGGATAGCGTTTACCCTGTTTAACTTGCTGACACCTGCTT	1140
Qy	1141	GGCGGCTTCGCGACAGAATTTGATTTCCGTCCGCTGGTGGCCAGCTGTTTCTACTCCCGTCC	1200
Db	1141	GGCGGCTTACCGACAGAATTTGATTTCCGTCCGCTGGTGGCCAGCTGTTTCTACTCCGTCC	1200
Qy	1201	GTGTCTCAGCAATGGCGAGCCGACTGTTTAAGTTGTATATCATCTGTAGAGATGCTCAG	1260
Db	1201	GTGTCTCAGCAATGGCGAGCCGACTGTTTAAGCTTTATACATCTGTAGAGATGCTCAG	1260

Qy	1261	CAGGATGAAGGGTATTGCAATCCCGCATGACATTCAGCTCGGAGAACTCTCGTGTGGTTATT	1320
Db	1261	CAGGATAAGGGTATTGCAATCCCGCATGACATCGACTCGGGGAGTCTCGTGTAGTTATT	1320
Qy	1321	CAGGATTATCATAAACCAACATGAACAGATCGCGCAGCGCTTCTCAGGCCCATCGCGC	1380
Db	1321	CAGGATTATGACAAACCAACATGAGAGGACCGACCGACACTTCCCGAGCCCATCGCGC	1380
Qy	1381	CCTTTCTCTCTCCCTTCGAGCTAATGATGTGCTTTTGGCTCTCTCTACCCGCTGCGAGTAT	1440
Db	1381	CCTTTTTCTGTCTCCGAGCTAATGATGTGCTTTTGGCTTCTCTCAACCGCTGCGGAGTAT	1440
Qy	1441	GACAGTCCACTATTAGGCTCTTGACATGGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG	1500
Db	1441	GACCAGTCCACTTACGSGCTCTTGACCGGCCACGTCATGTGCTCTGACTCTGTGACCTTG	1500
Qy	1501	GTTTAATGTTGCGACCGCGCGCAGCGCGTTGCCCGTGCCTCGATTGGACCAAGGTCACA	1560
Db	1501	GTTTAATGTTGGACCGCGCGCGCAGCGCGTTGCCCGGTCACTCGACTGGACCAAGGTCACA	1560
Qy	1561	CTTGACGCTCGCCCCCTCTCCACCATCCAGAGTACTCGAAGACCTTCTTTGCTCTGCCG	1620
Db	1561	CTTGATGGTCGCCCCCTTTCCACCATCCAGCAGTATTCAAAGACCTTCTTTGCTCTGCCG	1620
Qy	1621	CTCCGCGGTAAAGCTCTCTTTCTGGGAGCGAGGCACAACTAAAGCGGTACTCTATAAT	1680
Db	1621	CTCCGCGGTAAAGCTCTCTTTCTGGGAGCGAGGTACTATAAGCGGGTACCCCTTTAAT	1680
Qy	1681	TATAACACCACTGTACGCAACCACTCTTTGTCAGAAATCGCGCGGGCACCCGGGTGCT	1740
Db	1681	TATAACACCACTGCTAGTGACCACTGCTGTTGAGAAATCGCGCTGGGCATCGGTTGCT	1740
Qy	1741	ATTTCACCTTACACCACTAGCCTGGGTGCTGGTCCCGTCTCCAATTTCTGGGTTGCGGTT	1800
Db	1741	ATTTCACCTTACACCACTAGCCTGGGTGCTGGTCCCGTCTCTATTTCCGGGTGCTGTT	1800
Qy	1801	TTAGCCCCCACTCTGGGCTTAGCATTTGCTTTGAGAGATACCTTGGACTACCTGCCCGCGCC	1860
Db	1801	TTAGCCCCCACTCCGGGCTAGCATTTGCTTTGAGGATACCATGGACTACCTGCCCGCGCC	1860
Qy	1861	CATACTTTTGATGATTTCTGCCACAGAGTCGCCGCCCTTGCGCTTTCAGGGCTGCGCTTTC	1920
Db	1861	CATACTTTTCGATGACTTCTTGCCCGGAGTGGCCGCCCTTGGCCTCCAGGGCTGTGCTTTT	1920
Qy	1921	CAGTCTACTGTGCGTGAAGCTTTCAGCGGCTTAAAGATGAAGTGGGTGGAATAACTTCGGAGTTG	1980
Db	1921	CAGTCTACTGTGCGTGAAGCTTTCAGCGGCTTAAAGATGAAGTGGGTGGAATAACTTCGGAGTTA	1980
Qy	1981	TAG 1983	
Db	1981	TAG 1983	

RESULT 15
ABL60053
ID ABL60053 standard; DNA; 1990 BP.

XX ABL60053

16-AUG-2002 (first entry)

XX
DE Hepatitis E virus ORF-2 DNA #1.XX
KW
Hepatitis E virus: ORF-2: HEV:XX
05
Hepatitis E virus.XX
DN W0200240681-A1XX
DD 22-MAY-2002XX
20 SEP 2007

```
XX 30-SEP-2000; 2000CN-00130634.
XX (YANG-) YANG SHENG TANG CO LTD.
XX Xia N, Zhang J, Li S, Ge S, Gu Y, He Z;
XX WPI; 2002-427096/45.
XX Polypeptide fragments of hepatitis E virus (HEV) for vaccine compositions
XX and diagnosis in preventing and treating HEV infection.
XX Example 1; Page 82-83; 91pp; Chinese.
XX The sequence encodes hepatitis E virus (HEV) ORF-2 polypeptide. The
XX invention relates to novel n-mer polypeptides comprising amino acids of
XX ORF-2 of HEV where n = 2-180. The polypeptides of the invention have
XX virucide activity. The polypeptide of HEV, its fragments, and chimeric
XX proteins, are useful for vaccine compositions and diagnosis in preventing
XX and treating HEV infection by immunisation, which may also be applied in
XX the diagnosis of HEV infection
XX
XX SQ Sequence 1990 BP; 348 A; 666 C; 481 G; 495 T; 0 U; 0 Other;
XX
XX Query Match 91.0%; Score 1805.4; DB 6; Length 1990;
XX Best Local Similarity 94.4%; Pred. No. 0;
XX Matches 1872; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
XX
XX QY 1 ATGCGCCCTCGGCCCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGGCA 60
XX DB 1 ATGCGCCCTCGGCCCTATTTTGTGCTGCTCCTCATGTTTGTGCTATGCTGCGCGGCA 60
XX
XX QY 61 CGCGCCGCTCAGCGCTCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGT 120
XX DB 61 CGCGCCGCTCAGCGCTCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGT 120
XX
XX QY 121 TTCTGGGGTGACGGGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
XX DB 121 TTCTGGGGTGACGGGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
XX
XX QY 181 CCCTTGCAGCGGTGATGACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
XX DB 181 CCCTTGCAGCGGTGATGACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
XX
XX QY 241 CGACCACTCGGCTCGGCTGGCGGTGACAGCGCCAGCGCGCGCGCGCGCGCGCG 300
XX DB 241 CGACCACTCGGCTCGGCTGGCGGTGACAGCGCCAGCGCGCGCGCGCGCGCGCG 300
XX
XX QY 301 AGACCTACACAGCTGGGCGCGCGCGCTAACCGCGGTGCTCCGCGCCCATGACACCCCG 360
XX DB 301 AGACCTACACAGCTGGGCGCGCGCGCTAACCGCGGTGCTCCGCGCCCATGACACCCCG 360
XX
XX QY 361 CCAGTGCCTGATGCTGACTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
XX DB 361 CCAGTGCCTGATGCTGACTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
XX
XX QY 421 TCTCCCTTACCTTCTCGTGGCCACCGGTGACAGCGCGCGCGCGCGCGCGCGCG 480
XX DB 421 TCTCCCTTACCTTCTCGTGGCCACCGGTGACAGCGCGCGCGCGCGCGCGCGCG 480
XX
XX QY 481 AGTCCGCTTTTACCCCTTCAGGACGGCGCAATACCCCATATATGCGCCAGGAGCTTCT 540
XX DB 481 AGCCCACTTCTACCCCTTCAGGACGGCGCAATACCCCATATATGCGCCAGGAGCTTCT 540
XX
XX QY 541 AATTATGCCAGTACGGGTTCGCGTGCACCAATTCGGTACCGCGCGGTGTCGCCAAT 600
XX DB 541 AATTATGCCAGTACGGGTTCGCGTGCACCAATTCGGTACCGCGCGGTGTCGCCAAT 600
XX
XX QY 601 GCTGTGCGGGTTACGCCATCTCCATCTCATTTCTGCGCACAGACCAACACCCCGGAG 660
XX DB 601 GCTGTGCGGGTTACGCCATCTCCATCTCATTTCTGCGCACAGACCAACACCCCGGAG 660
XX
XX QY 661 TCCGTTGATATGAATTCAATACCTCGACGGATGTTTCGTTATTTTAGTCCAGCCCGGCGATA 720
XX DB 661 TCCGTTGATATGAATTCAATACCTCGACGGATGTTTCGTTATTTAGTCCAGCCCGGCGATA 720
XX
XX 661 TCCGTTGATATGAATTCAATACCTCGACGGATGTTTCGTTATTTAGTCCAGCCCGGCGATA 720
XX QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATCTATCTAACCAAGGCTGCGCTCC 780
XX DB 721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATCTATCTAACCAAGGCTGCGCTCC 780
XX QY 781 GTGAGACCTCTGGGCTGAGGAGGAGGCTACCTCTGCTCTGTTTATGCTTGTGCTGATA 840
XX DB 781 GTGAGACCTCTGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
XX QY 841 CATGGCTCACTCGTAAATTCCTATATCTATATACCTTATACCTTATACCTTATACCTTATACCTT 900
XX DB 841 CATGGCTCACTCGTAAATTCCTATATCTATATACCTTATACCTTATACCTTATACCTTATACCTT 900
XX QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAACCAACCAACCAACCAAC 960
XX DB 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAACCAACCAACCAACCAAC 960
XX QY 961 TCCGTTATTCAGGACTGCTGCGCACCGCTTCTGCGCGGTGCGGAGCGGAGCTGCGGAG 1020
XX DB 961 TCCGTTATTCAGGACTGCTGCGCACCGCTTCTGCGCGGTGCGGAGCGGAGCTGCGGAG 1020
XX QY 1021 CTCACCAACCAACCGCTGCTACCCGCTTTATGAAGGAGCTTCTATTTTACTAGTACTATAGT 1080
XX DB 1021 CTCACCAACCAACCGCTGCTACCCGCTTTATGAAGGAGCTTCTATTTTACTAGTACTATAGT 1080
XX QY 1081 GTCGTTGAGATCGGCGCGGATAGCCCTCACCCCTGTTCAACCTGCTGACACTCTGCTT 1140
XX DB 1081 GTCGTTGAGATCGGCGCGGATAGCCCTCACCCCTGTTCAACCTGCTGACACTCTGCTT 1140
XX QY 1141 GCGGCGCTGCGGAGAGATTTGCTGCGGCTGCTGCGGAGCTGCTTCTACCTCCGCTCC 1200
XX DB 1141 GCGGCGCTGCGGAGAGATTTGCTGCGGCTGCTGCGGAGCTGCTTCTACCTCCGCTCC 1200
XX QY 1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGTATACATCTCTAGAGAACTGCTAG 1260
XX DB 1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGTATACATCTCTAGAGAACTGCTAG 1260
XX QY 1261 CAGGATTAAGGTTATGCAATCCCGATGACATCGACCTCGGCGGAGTCTCGTGTAGTTAT 1320
XX DB 1261 CAGGATTAAGGTTATGCAATCCCGATGACATCGACCTCGGCGGAGTCTCGTGTAGTTAT 1320
XX QY 1321 CAGGATTAAGTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1380
XX DB 1321 CAGGATTAAGTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1380
XX QY 1381 CCTTTCTCTGCTTCTGAGCTAATGATGCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
XX DB 1381 CCTTTCTCTGCTTCTGAGCTAATGATGCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
XX QY 1441 GACCACTCACTTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
XX DB 1441 GACCACTCACTTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
XX QY 1501 GTTAAATGTTGCGAGCGGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
XX DB 1501 GTTAAATGTTGCGAGCGGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
XX QY 1561 CTTGACCGGTGCGCCCTCTCTCAACCACTCAGAGAGTACTCGAGAGCTTCTTTGCTGCTGCGG 1620
XX DB 1561 CTTGATGCTGCGCCCTCTCTCAACCACTCAGAGAGTACTCGAGAGCTTCTTTGCTGCTGCGG 1620
XX QY 1621 CTCGCGGTAAAGTCTCTCTTTTGGGAGGAGGAGTACTTAAAGCGGGTACCTTATAT 1680
XX DB 1621 CTCGCGGTAAAGTCTCTCTTTTGGGAGGAGGAGTACTTAAAGCGGGTACCTTATAT 1680
XX QY 1681 TATAACCACTGCTAGGAGCAACTGCTTGTGAGAAATGCGGCGGAGCGGAGCTGCTGCTGCTGCTGCTGCT 1740
XX DB 1681 TATAACCACTGCTAGGAGCAACTGCTTGTGAGAAATGCGGCGGAGCGGAGCTGCTGCTGCTGCTGCTGCT 1740
XX QY 1741 ATTTCCACTTACCACTAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
XX DB 1741 ATTTCCACTTACCACTAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
```

Db 1741 ATTTCACATTACACCACTAGCCTGGGTGCTGGTCCCGTCTCTATTTCGGCGTTGCTGTT 1800
QY 1801 TTAGCCCCCACACTCTGCGCTAGCATTTGCTTTGAGGATACCTTTGGACTACCTTGCCCGCGCC 1860
Db 1801 TTAGCCCCCACACTCCGCGCTAGCATTTGCTTTGAGGATACCATGGACTACCTTGCCCGCGCC 1860
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTGCCGCCCTTGCGCTTCAGGGCTGGCTTTC 1920
Db 1861 CATACTTTCGATGACTTCTGCCGAGTGCCGCCCTTGCGCTTCAGGGCTGGCTTTT 1920
QY 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGATGAAGTGGTAAACTCGGGAGTTG 1980
Db 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGATGAAGTGGTAAACTCGGGAGTTA 1980
QY 1981 TAG 1983
Db 1981 TAG 1983

Search completed: August 21, 2004, 11:00:11
Job time : 775.579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 10:14:02 ; Search time 5116.87 seconds
(without alignments)
11572.836 Million cell updates/sec

Title: US-09-851-410a-6_COPY_5147_7129

Perfect score: 1983

Sequence: 1 ATGCGCCCTCGGCTATTTT.....GTAAACTCGGGATTGTAG 1983

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmi:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67.2	3.4	925	29	CNS0091P
C 2	67.2	3.4	1201	9	AL578906
C 3	63.6	3.2	1201	13	AL578906
C 4	62.2	3.1	1103	13	AL578906

C 5	60.2	3.0	1201	13	BX360624
C 6	60.2	3.0	1839	29	CG747711
C 7	60	3.0	1057	13	BX349688
C 8	59.6	3.0	844	29	CNS0052P
C 9	59.6	3.0	1201	13	BX446579
C 10	59.2	3.0	839	29	CNS004NB
C 11	59	3.0	1793	29	CG754612
C 12	58.4	2.9	776	29	CNS010RY
C 13	58	2.9	932	29	CNS0072Q
C 14	58	2.9	1174	14	CD508311
C 15	58	2.9	1213	12	BM560246
C 16	57.6	2.9	1695	28	CC290874
C 17	57.4	2.9	1135	28	BZ574895
C 18	57.4	2.9	1970	29	CG748837
C 19	57.2	2.9	807	13	BX464554
C 20	57.2	2.9	935	29	CNS006XK
C 21	57	2.9	1136	14	CD507542
C 22	56.8	2.9	717	29	AG135089
C 23	56.4	2.8	946	29	AG080623
C 24	56.4	2.8	1000	13	BX407619
C 25	56.4	2.8	1061	13	BX391246
C 26	55.8	2.8	1007	28	BZ569268
C 27	55.8	2.8	1157	28	CC219063
C 28	55.6	2.8	964	29	CNS003WG
C 29	55.6	2.8	1016	13	BQ681552
C 30	55.6	2.8	1094	14	CD509997
C 31	55.6	2.8	1193	13	BX421125
C 32	55.4	2.8	1137	29	AG078502
C 33	55.4	2.8	1686	28	CC189908
C 34	55.2	2.8	995	14	CA788818
C 35	55.2	2.8	1032	13	BQ680891
C 36	55.2	2.8	1356	29	CG754548
C 37	54.8	2.8	1362	12	BM556092
C 38	54.6	2.8	948	13	BQ53470
C 39	54.6	2.8	1107	28	BZ051753
C 40	54.6	2.8	1167	13	BQ061197
C 41	54.4	2.7	703	29	AG167751
C 42	54.4	2.7	739	29	AG080613
C 43	54.4	2.7	982	13	BX415111
C 44	54.4	2.7	983	13	BX38409
C 45	54.4	2.7	1036	9	AL550172

ALIGNMENTS

RESULT 1	CNS0091P	925 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
DEFINITION	BAC19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL053013				
VERSION	AL053013.1				
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Ephyridiacea; Endopterygota; Diptera; Brachycera; Muscomorpha; 1 (bases 1 to 925)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseogawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of				

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila/bac.htm>.

FEATURES

```
1. 925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"
```

ORIGIN

Query Match	3.4%;	Score 67.2;	DB 29;	Length 925;
Best Local Similarity	12.4%;	Pred. No.0.0001;		
Matches	44;	Conservative 180;	Mismatch 132;	Indels 0; Gaps 0;
50	TGCCCGGCACCGCCCGGTGACCGCTCTGGCGCGCGCTCTGGCGGCGGACGACGCGCGTT	109		
570	BSCGCSGCSGSSSCCBCCGCCSCSYSCSSBBSBKCESTBSBSCCCCSKSVCGT	629		
110	CGGGCGGTGTTCTCGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTC	169		
630	SCSSSSCSGSSSSTSSSTSSSTSKSSGSSGSSSSSYTTSKTSASGSGSNWSAGGS	689		
170	ATCCAAACCAACCCCTTGCGCCCGATGTACACCGCTGGCGCGGGCTGGACTGTGTTC	229		
690	GSTGTSTSSSSSSTSTSSSSVSQSKSTBSGSSBSGSSSSSTSSBBSCTSTSSS	749		
230	GCCAAACCGCGCCGACCACTCGGTCCGCTTGCGCTGACCAAGCCACCGCCCGCCGCTG	289		
750	SSSYSSSTGSCCTCCCSYSYSSSTSSSSTWSGTSGSSSSSVGTSSSSDSTSTCCSCC	809		
290	CCTCAGCTGTAGACCTTACCACAGCTGGGGCGCGCGCTAACCGCGGTGCGTCCGGCCC	349		
810	YMCCTCSTYMBCYTSTSCGSSSSSGKGGVTKCGCGGCGSSSTNGMBTSSACSSSSS	869		
350	ATGACACCCGCGAGTGCCTGATGTGCACTCCCGCGCGCCCACTTGGCGCGGAC	405		
870	CSSSVSSSSKSGAASSSVSSGCGSGVSNNSGNSAKSKSSSGSVSSGSGGSGSVG	925		

RESULT 2

AL578906 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK005YA19 5-PRIME, mRNA sequence.

1201 bp mRNA linear EST 01-JUN-2003

AL578906 Homo sapiens (human)

AL578906 Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1. (Bases 1 to 1201), Jesse, J. and Polayes, D.
Li, W.B., Gruber, C.,
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:12943429.

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 584.f For
more information about this cluster, see
<http://www.genoscope.cns.fr/>
<http://bin.cluster.cgi?seq=CS0DK005AA10QPI&cluster=584.f>. Contact :

Peng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> InvitroGen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DK005AA10QP1.

FEATURES

```

1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cloned="CS0DK005YA19"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/cloned_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="First strand cDNA was primed with a
primer. Five prime end enriched, double-stranded
cDNA was digested with Not I and cloned into the Not I
sites of the pCMVSPORT6 vector. Library
size: 1.2 x 10^6"

```

ORIGIN

[illegible]

RESULT 3

RESIDU_3	BX335523/c			1201 bp	mRNA	linear	EST 01-MAY-2000
LOCUS	BX335523	Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSODI01SYN12 3-PRIME, mRNA sequence.					
DEFINITION							
ACCESSION	BX335523						
VERSION	1	GI:30312422					
KEYWORDS	EST.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
REFERENCE							
AUTHORS	Li, W.B., Gruber,C., Jessee,J. and Polayes,D.						
TITLE	Full-length cDNA libraries and normalization						
JOURNAL	Unpublished (2001)						
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen.FEATURES Location/Qualifiers Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen corporation 1600 Faraday Avenue Genoscope sequence ID : CSODI015DG06NP1. Location/Qualifiers						
FEATURES							

/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="CS0D1072YL05"	
/issue_type="PLACENTA COT 25-NORMALIZED"	
/clone_lib="Homo sapiens PLACENTA COT 25-normalized"	
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."	
ORIGIN	
Query Match 3.0%; Score 60.2; DB 13; Length 1201;	
Best Local Similarity 9.2%; Pred. No. 0.0052;	
Matches 25; Conservative 163; Mismatches 85; Indels 0; Gaps 0;	
QY	65 CCGGTACGCGTCTGCGCGCGCGTGTGGCGGCGACGCGCGTTCGGCGGTTCCT 124
DB	1178 SBTSTTTSTBTSSSSSSSBSTTTSSSSSBSTSSSSSBSTSSSSSBSTSS 1119
QY	125 GGGGTACCGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACACCCCT 184
DB	1118 TTTTTSVTTCTBTWSKTTSTCSYTTCTVTSYCSBECSTTSCCTTCTCTCV 1059
QY	185 TCGCCCCCGATGCACGCTCGCGCGCGCGCTCGACCTCTGTTTCGCCAACCCCGCGAC 244
DB	1058 STSSBCCSBTTTSSSSSSSSSTSTSTSSSSSSSSSSSSSSSSSSSSSS 999
QY	245 CACTCGCTCCGTTGCGGTGACAGCCGCGCGCGCGCGCGTTCGCTACGTCGTAGAC 304
DB	998 SSSSSBTTSSSSSSSBSTSSSSSSSSSSSSSSSSSSSSSSSBSTSSSSSS 939
QY	305 CTACACAGTGGGCGCGCGCTAACCGCG 337
DB	938 SSSSSSSSSSSSSSSSTSKWSSTASSSSSV 906
RESULT 6	
LOCUS	CG747711/c 1839 bp DNA linear GSS 24-OCT-2003
DEFINITION	P041-2-E02.2a Ppa EcoRI BAC Library Pristionchus pacificus genomic, genomic survey sequence.
ACCESSION	CG747711
VERSION	CG747711.1 GI:37968637
KEYWORDS	GSS.
SOURCE	Pristionchus pacificus
ORGANISM	Pristionchus pacificus
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
AUTHORS	Neodiplogasteridae; Pristionchus.
TITLE	1 (bases 1 to 1839)
JOURNAL	Srinivasan, J., Sinz, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K.,
MEDLINE	Buntjer, J., van der Meulen, M. and Sommer, R.J.
PUBMED	An integrated physical and genetic map of the nematode Pristionchus pacificus
COMMENT	Mol. Genet. Genomics 269 (5), 715-722 (2003)
Contact: Sommer RJ	
Evolutionary Biology	
Max-Planck-Institute for Developmental Biology	
Spemannstr. 37-39, Tuebingen D-72076, Germany	
Tel: 00497071601371	
Fax: 00497071601498	
Email: ralf.sommer@tuebingen.mpg.de	
Class: BAC ends.	
Location/Qualifiers	
1..1839	
/organism="Pristionchus pacificus"	
/mol_type="genomic DNA"	
/strain="California"	
/db_xref="taxon:54126"	
/clone_lib="Ppa EcoRI BAC Library"	
/note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC	
FEATURES	
source	
ORIGIN	
Query Match 3.0%; Score 60.2; DB 29; Length 1839;	
Best Local Similarity 37.2%; Pred. No. 0.0064;	
Matches 194; Conservative 0; Mismatches 327; Indels 0; Gaps 0;	
QY	143 CTCAGCCCTTGGCAATCCCTATATTCATCCAAACACCCCTTCGCGCCGATGTACCG 202
DB	1368 CCC 1309
QY	203 CTCGCGCGGGGTGGACCTGTGTTTGGCAACCCCGCGACCACTCGGCTTCGGTGGC 262
DB	1308 CCC 1249
QY	263 GTGACAGGCGCCAGCGCGCGTTCCTCAGTCGTAGACTACACAGCTGGGCGCG 322
DB	1248 CCNNCC 1189
QY	323 CGCGCGCTAACCGCGGTGCTTCGGCGCCCATGACACCCCGCAGTGCCTGATGTGCACTCC 382
DB	1188 CCCCCNNNNCCCCNNNNCCCCNNCCCCCCCCCCCCCCCCCCCCCCCCCCCC 1129
QY	383 GCGGCGCATCTTGGCGCGGAGTATAAATATCAACATCTCCCTTACCTTTCGTGG 442
DB	1128 CCC 1069
QY	443 CCACCGGCCTAACCTGGTCTTTATGCGCGCCCTCTTAGTCGCTTTTACCCCTTCAGG 502
DB	1068 CCCCCNNNNCCCCNNNNCCCCNNCCCCCCCCCCCCCCCCCCCCCCCCCCCC 1009
QY	503 ACGGCACCAATACCATATAATGCGCACGGAAGTCTTAATTATGCCCAGTACCGGGTTG 562
DB	1008 CCNCCCCCCCCCCCCCCCCCCCCCCCCNNNNCCNNCCNNCCNNCCNNCCNNNN 949
QY	563 CCGGTGCCACAATCGGTTACCGCGGTGTGTCGCAATGCTGTGCGGGTTACGCGTCT 622
DB	948 CCCCCNNNNCCCCNNNNCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 889
QY	623 CCATCTCATTTCTGGCCACAGACACCAACCCAGCGAGTCC 663
DB	888 CCNCCCCCCCCCCCCCCCCCCCCNNCCNNCCCCCCCCCCCCCCCCCCCC 848
RESULT 7	
LOCUS	BX349688/c 1057 bp mRNA linear EST 05-MAY-2003
DEFINITION	BX349688 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0D1063YK05 3-PRIME, mRNA sequence.
ACCESSION	BX349688
VERSION	BX349688.1 GI:30375394
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1057)
TITLE	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
JOURNAL	Full-length cDNA libraries and normalization
COMMENT	Unpublished (2001)
Contact: Genoscope	
Genoscope - Centre National de Sequencage	
BP 191 91006 EVRY cedex - France	
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr	
Library was constructed by Life Technologies, a division of	
Invitrogen. This sequence belongs to sequence cluster 7092.r For	
more information about this cluster, see	
http://www.genoscope.cns.fr/	
cgi-bin/cluster.cgi?seq=CS0BAH015ZB03NP1&cluster=7092.r. Contact :	
Feng Liang Email : fliang@lifetech.com URL :	
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600	
Faraday Avenue Genoscope sequence ID : CS0BAH015ZB03NP1.	
Location/Qualifiers	
FEATURES	

```

source
1..1057
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01063YK05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 3.0%; Score 60; DB 13; Length 1057;
Best Local Similarity 32.1%; Pred. No. 0.0055;
Matches 122; Conservative 77; Mismatches 181; Indels 0; Gaps 0;

QY 150 CTTGCAATCCCTATATTCATCAACCAACCCCTTGCAGCCGAGTGCACGCTGCGC 209
D 1111111111111111111111111111111111111111111111111111111
692 CATCGCTATTAACCATGTGATCGGCTCTCGCGCCGCCCCCCYGGCGCGCCGCGC 633
D 1111111111111111111111111111111111111111111111111111111
QY 210 CGGGCTGACCTGTTTCGCAACCCGCCGACCTCGCTCGCTCGCTCGCTCGCTCA 269
D 1111111111111111111111111111111111111111111111111111111
632 GCGCCSSCGSCGSGSCSCGCGCCGCTCTCCSCGCGCCGCGCCGCCGCCGCCGCC 573
D 1111111111111111111111111111111111111111111111111111111
QY 270 GCGCCAGCGCCCGCGCTTCTCCTCATCTAGCTAGCTACACAGCTGGCGCGCGCT 329
D 1111111111111111111111111111111111111111111111111111111
572 CCCCCSSCGSSCGSSCGSSCGSSCGSSCGSSCGSSCGSSCGSSCGSSCGSSCGSS 513
D 1111111111111111111111111111111111111111111111111111111
QY 330 AACCGGCTGCTCGGCGCATGACACCGCGCAGTGCCTGATGTCGATCGCGCGGCG 389
D 1111111111111111111111111111111111111111111111111111111
512 SCGSSSSSGSSCGCGCGSCGCGSCGCGSCGCGSCGCGSCGCGSCGCGSCGCGSCG 453
D 1111111111111111111111111111111111111111111111111111111
QY 390 CATCTTGGCGCGCAGTATAACCTATCAATCTCCCTTACTCTTCGTCGCGCACCG 449
D 1111111111111111111111111111111111111111111111111111111
452 CCGCCGCGSCGSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 393
D 1111111111111111111111111111111111111111111111111111111
QY 450 CACTAACCTGCTTTTATCGCGCGCGCTTTAGTCGCTTTTACCCCTTCAGGACGG 509
D 1111111111111111111111111111111111111111111111111111111
392 CGCCCCSSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 333
D 1111111111111111111111111111111111111111111111111111111
QY 510 CAATACCATATATGGCA 529
D 1111111111111111111111111111111111111111111111111111111
332 MSCNTMSMCCATNAGGCGCTA 313
D 1111111111111111111111111111111111111111111111111111111

RESULT 8
CNS0052P/c
LOCUS
DEFINITION
CNS0052P 844 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR11P16 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL056652
AL056652.1 GI:4932342
GSS.
Drosophila melanogaster (fruit fly)
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 844)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of

```

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1..844
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR11P16"
/clone_lib="RPCL-98"
/note="end : TET3"

ORIGIN

```

Query Match 3.0%; Score 59.6; DB 29; Length 844;
Best Local Similarity 19.9%; Pred. No. 0.0061;
Matches 59; Conservative 127; Mismatches 111; Indels 0; Gaps 0;

QY 21 GTTGCTGCTCTCATGTTTTCCTATGCTGCGCGCCACCGCGCTCAGCGCTCGG 80
D 1111111111111111111111111111111111111111111111111111111
678 BTTCWTHNSNTYTWTVYVTHHTYVTHHTYVTHHTYVTHHTYVTHHTYVTHHTY 619
D 1111111111111111111111111111111111111111111111111111111
QY 81 CGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 140
D 1111111111111111111111111111111111111111111111111111111
618 CCGCGSSSSSSCGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSS 559
D 1111111111111111111111111111111111111111111111111111111
QY 141 TTCTCAGCGCTTCGCAATCCCTATATTCATCAACCAACCCCTTCCGCGCGAT 200
D 1111111111111111111111111111111111111111111111111111111
558 CKCBSSSSSSBSSBSSSSCCYCSYCSYCSYCSYCSYCSYCSYCSYCSYCSYCSY 499
D 1111111111111111111111111111111111111111111111111111111
QY 201 CGCTCGCGCGCGCGCTGAGCTCTGTTTCGCAACCCCGCGCGCGCGCGCGCT 260
D 1111111111111111111111111111111111111111111111111111111
498 SGGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSG 439
D 1111111111111111111111111111111111111111111111111111111
QY 261 GGTGACACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 317
D 1111111111111111111111111111111111111111111111111111111
438 SBBGGSSSKTSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSS 382
D 1111111111111111111111111111111111111111111111111111111

```

RESULT 9

BX446579 1201 bp mRNA linear EST 22-MAY-2003
BX446579 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CL0BB002ZF10 3-PRIME, mRNA sequence.

ACCESSION
BX446579.1 GI:31033746

KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1201)

AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7325.f For
more information about this cluster, see

<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CL0BB002ZF10&cluster=7325.f>. Contact :
Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>

Faraday Avenue Genoscope sequence ID : CL0BB002ZF10PPI.

FEATURES

Location/Qualifiers

1..1201

source

Blank Sheet

QY 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCCCGGCA 60
DB 1 ATGCGCCCTCGGCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCCCGGCA 60
QY 61 CGCGCCGGTACGCGTCTGCGCGCGCTGCTGCGCGCGCGAGCGGGTTCGCGCGTGT 120
DB 61 CGCGCCGGTACGCGTCTGCGCGCGCTGCTGCGCGCGCGAGCGGGTTCGCGCGTGT 120
QY 121 TTCTGCGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
DB 121 TTCTGCGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
QY 181 CCTTTCGCCCGCGATGTCAGCGTCTGCGCGCGCGGGCTGAGCTGCTGTCGCCAAACCGGC 240
DB 181 CCTTTCGCCCGCGATGTCAGCGTCTGCGCGCGCGGGCTGAGCTGCTGTCGCCAAACCGGC 240
QY 241 CGACCACTCGGCTCGCTTGGCGTGACACGCGCCAGCGCCCGCGCTTCCTCAGTCTGT 300
DB 241 CGACCACTCGGCTCGCTTGGCGTGACACGCGCCAGCGCCCGCGCTTCCTCAGTCTGT 300
QY 301 AGACCTACACAGTGGGCGCGCGCTTAACCGCGGTGCTTCGCGCCCATGACACCCCG 360
DB 301 AGACCTACACAGTGGGCGCGCGCTTAACCGCGGTGCTTCGCGCCCATGACACCCCG 360
QY 361 CGAGTGCCTGATGTCAGCTCCCGCGCGCATCTTCGCGCGCGAGTATTAACCTATCAACA 420
DB 361 CGAGTGCCTGATGTCAGCTCCCGCGCGCATCTTCGCGCGCGAGTATTAACCTATCAACA 420
QY 421 TCTCCCTTACCTCTTCCGTTGGCCACCGGCACTTAACCTGCTTCTTTATGCGCCCTCTT 480
DB 421 TCTCCCTTACCTCTTCCGTTGGCCACCGGCACTTAACCTGCTTCTTTATGCGCCCTCTT 480
QY 481 AGTCCGCTTTACCTCTTCCGTTGGCCACCGGCACTTAACCTGCTTCTTTATGCGCCCTCTT 540
DB 481 AGTCCGCTTTACCTCTTCCGTTGGCCACCGGCACTTAACCTGCTTCTTTATGCGCCCTCTT 540
QY 541 AATTATGCCAGTACCGGGTTCGCGTGCACAACTCCGTTACCGCCCGCTGCTGCCAAT 600
DB 541 AATTATGCCAGTACCGGGTTCGCGTGCACAACTCCGTTACCGCCCGCTGCTGCCAAT 600
QY 601 GCTGTCCGGGTTACGCCATCTCCATCTATTCTGGCCACAGACCAACCCCGAGC 660
DB 601 GCTGTCCGGGTTACGCCATCTCCATCTATTCTGGCCACAGACCAACCCCGAGC 660
QY 661 TCGGTTGATGATGAATCAATAAAGTCTGAGAGGAGGCTACCTCTGGTCTGTTATGCTTGCATA 720
DB 661 TCGGTTGATGATGAATCAATAAAGTCTGAGAGGAGGCTACCTCTGGTCTGTTATGCTTGCATA 720
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGTCC 780
DB 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGTCC 780
QY 781 GTCGAGACCTCTCGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTGTTATGCTTGCATA 840
DB 781 GTCGAGACCTCTCGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTGTTATGCTTGCATA 840
QY 841 CATGGCTCACTGTAATTCCTATATACTAATACACCTATACCGTGCCTCGCGGCTGTTG 900
DB 841 CATGGCTCACTGTAATTCCTATATACTAATACACCTATACCGTGCCTCGCGGCTGTTG 900
QY 901 GACTTTGCGCTTCGAGCTTGTGATTTGCAACCTTACCCCGGTAAACCAATAACCGGGTC 960
DB 901 GACTTTGCGCTTCGAGCTTGTGATTTGCAACCTTACCCCGGTAAACCAATAACCGGGTC 960
QY 961 TCCCGTTATTCAGCACTGCTCGCACCGCTTGTGTCGGGTGCGGACGGAGTGGCGAG 1020
DB 961 TCCCGTTATTCAGCACTGCTCGCACCGCTTGTGTCGGGTGCGGACGGAGTGGCGAG 1020
QY 1021 CTCACACACCGGTGCTACCGGCTTTATGAAGACCTCTATTTTACTAGTACTATAGT 1080
DB 1021 CTCACACACCGGTGCTACCGGCTTTATGAAGACCTCTATTTTACTAGTACTATAGT 1080
QY 1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1140

DB 1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1140
QY 1141 GCGCGCTCGCGACAGAAATTGATTTTCGCGGTGCTGGCGAGCTGTTCTACTCCGCTCC 1200
DB 1141 GCGCGCTCGCGACAGAAATTGATTTTCGCGGTGCTGGCGAGCTGTTCTACTCCGCTCC 1200
QY 1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
DB 1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
QY 1261 CAGGATAAGGGTATTGCAATCCCGATGACATTGACCTCGGAGAAATCTCGTGTGGTTAT 1320
DB 1261 CAGGATAAGGGTATTGCAATCCCGATGACATTGACCTCGGAGAAATCTCGTGTGGTTAT 1320
QY 1321 CAGGATAAGGGTATTGCAATCCCGATGACATTGACCTCGGAGAAATCTCGTGTGGTTAT 1380
DB 1321 CAGGATAAGGGTATTGCAATCCCGATGACATTGACCTCGGAGAAATCTCGTGTGGTTAT 1380
QY 1381 CTTTCTCTGCTTTCGAGCTAATGATGCTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
DB 1381 CTTTCTCTGCTTTCGAGCTAATGATGCTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
QY 1441 GACCACTCCACTTATGGCTCTTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
DB 1441 GACCACTCCACTTATGGCTCTTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
QY 1501 GTTAATGTTGCAACCGCGCGGAGCGCTTGGCGGTGCTCGATTTGGAACCAAGGTCA 1560
DB 1501 GTTAATGTTGCAACCGCGCGGAGCGCTTGGCGGTGCTCGATTTGGAACCAAGGTCA 1560
QY 1561 CTTGAGCGTGGCGCCCTCTCCACCATCCAGAGTACTCTGAAAGACCTTCTTGTCTGCGG 1620
DB 1561 CTTGAGCGTGGCGCCCTCTCCACCATCCAGAGTACTCTGAAAGACCTTCTTGTCTGCGG 1620
QY 1621 CTTCCGGGTAAAGCTCTCTTCTGGGAGGAGGACAACTAAAGCCGGGTACCTTATAAT 1680
DB 1621 CTTCCGGGTAAAGCTCTCTTCTGGGAGGAGGACAACTAAAGCCGGGTACCTTATAAT 1680
QY 1681 TATAACACCACTGCTAGACCAACTGTTGTCGAAATGCGCGCGGACCGGCTGCT 1740
DB 1681 TATAACACCACTGCTAGACCAACTGTTGTCGAAATGCGCGCGGACCGGCTGCT 1740
QY 1741 ATTTCCACTTACACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB 1741 ATTTCCACTTACACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
QY 1801 TTAGCCCCCACTCTCGCTTAGCAATGCTTGGAGTACCTTGGACTACCTGCTGCTGCTGCT 1860
DB 1801 TTAGCCCCCACTCTCGCTTAGCAATGCTTGGAGTACCTTGGACTACCTGCTGCTGCTGCT 1860
QY 1861 CATACTTTGATGATTTGCTGCGGAGTGGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCG 1920
DB 1861 CATACTTTGATGATTTGCTGCGGAGTGGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCG 1920
QY 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGTAAAGTGGGTAAACCTCGGAGTGG 1980
DB 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGTAAAGTGGGTAAACCTCGGAGTGG 1980
QY 1981 TAG 1983
DB 1981 TAG 1983

RESULT 2
US-08-259-148A-3
; Sequence 3, Application US/08259148A
; Patent No. 5741490
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, JF-Shin
; APPLICANT: Purdy, Michael A.

APPLICANT: Tam, Albert W.
 APPLICANT: Krawczynski, Krzysztof Z.
 APPLICANT: Yarbough, Patrice D.
 TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
 NUMBER OF SEQUENCES: 60
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 350 Cambridge Avenue, Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/259,148A

FILING DATE: 13-JUN-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 822,335

FILING DATE: 17-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 505,888

FILING DATE: 05-APR-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 420,921

FILING DATE: 13-OCT-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 367,486

FILING DATE: 16-JUN-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 336,672

FILING DATE: 11-APR-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 208,997

FILING DATE: 17-JUN-1988

ATTORNEY/AGENT INFORMATION:

NAME: Sholtz, Charles K.

REGISTRATION NUMBER: 38,615

REFERENCE/POCKET NUMBER: 4600-0093.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO. 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 2049 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 7

US-08-259-148A-3

Query Match 100.0%; Score 1983; DB 1; Length 2049;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCCCTCGCCCTATTGCTGCTCTCTCATGTTTTTTCCTATGCTGCGCGGCCA	60
Db	1	ATGGCCCTCGCCCTATTGCTGCTCTCTCATGTTTTTTCCTATGCTGCGCGGCCA	60
Qy	61	CCGCCCCGTACGCCGTCTGCGCGCGCGTCTGCGCGCGCGCGCGCGCGCGCGG	120
Db	61	CCGCCCCGTACGCCGTCTGCGCGCGCGTCTGCGCGCGCGCGCGCGCGCGCGG	120
Qy	121	TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC	180
Db	121	TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC	180

Qy	181	CCCTTCGCCCCGATGTGACCCCTGCGGCGCGGGGTGGACCTGCTGTTTCGCCAACCCGCC	240
Db	181	CCCTTCGCCCCGATGTGACCCCTGCGGCGCGGGGTGGACCTGCTGTTTCGCCAACCCGCC	240
Qy	241	CGACCACTCGGCTCGGCTTGGGCTGACACAGGCCCGCGCGCGCGCGCGCGCGCGCG	300
Db	241	CGACCACTCGGCTCGGCTTGGGCTGACACAGGCCCGCGCGCGCGCGCGCGCGCGCG	300
Qy	301	AGACCTACCAAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	360
Db	301	AGACCTACCAAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	360
Qy	361	CCAGTGCCCTGATGTCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	420
Db	361	CCAGTGCCCTGATGTCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	420
Qy	421	TCCTCCCTTACCTCTTCGCTGGCCACCGCGCACTAACCTGGTGTCTTTTATGCGCCCTCT	480
Db	421	TCCTCCCTTACCTCTTCGCTGGCCACCGCGCACTAACCTGGTGTCTTTTATGCGCCCTCT	480
Qy	481	AGTCCGCTTTTACCCCTTCAAGGACCGCACCAATACCCATATAATGCGCCACGGAAGCTT	540
Db	481	AGTCCGCTTTTACCCCTTCAAGGACCGCACCAATACCCATATAATGCGCCACGGAAGCTT	540
Qy	541	AATTATGCCAGTACCGGGTTGCCGCTGCGCAATCCGTTACCGCGCGCGTGTCCCAAT	600
Db	541	AATTATGCCAGTACCGGGTTGCCGCTGCGCAATCCGTTACCGCGCGCGTGTCCCAAT	600
Qy	601	GCTGTCGGCGGTTACGCCATCTCCATCTCATCTCTGGCCACAGACCACACCCCGGAG	660
Db	601	GCTGTCGGCGGTTACGCCATCTCCATCTCATCTCTGGCCACAGACCACACCCCGGAG	660
Qy	661	TCCGTTGATATGAATTCAATTAACCTCGACGATGTTGGTATTTAGTCCAGCCCGGATA	720
Db	661	TCCGTTGATATGAATTCAATTAACCTCGACGATGTTGGTATTTAGTCCAGCCCGGATA	720
Qy	721	GCCTCTGAGCTTGATCCCAAGTGAGGCGCTACACTATCGTAACCAAGGTGGCGCTCC	780
Db	721	GCCTCTGAGCTTGATCCCAAGTGAGGCGCTACACTATCGTAACCAAGGTGGCGCTCC	780
Qy	781	GTCGAGACTCTGCGGCTGAGGAGGAGGCTACCTCTGCTGTTGTTATGCTTGCATA	840
Db	781	GTCGAGACTCTGCGGCTGAGGAGGAGGCTACCTCTGCTGTTGTTATGCTTGCATA	840
Qy	841	CATGGCTCACTCGTAAATTCCTATATAACCCCTATACCCCTATACCGGTGCGCTGTTG	900
Db	841	CATGGCTCACTCGTAAATTCCTATATAACCCCTATACCCCTATACCGGTGCGCTGTTG	900
Qy	901	GACTTTCGCTTGGCTGAGCTTGATTCGCAACCTTACCCCGCGTAACCAATACCGGGTC	960
Db	901	GACTTTCGCTTGGCTGAGCTTGATTCGCAACCTTACCCCGCGTAACCAATACCGGGTC	960
Qy	961	TCCCGTTATTCAGCACTGCTCGCCACCGGCTTCGTCGCGTGGCGGAGGACGCGGAG	1020
Db	961	TCCCGTTATTCAGCACTGCTCGCCACCGGCTTCGTCGCGTGGCGGAGGACGCGGAG	1020
Qy	1021	CTACCAACCAAGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTAAATGGT	1080
Db	1021	CTACCAACCAAGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTAAATGGT	1080
Qy	1081	GTCCGTGAGATCGCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTT	1140
Db	1081	GTCCGTGAGATCGCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTT	1140
Qy	1141	GGCGGCTGCGGACAGAAATGATTTGCTGCGTGGTGGCGGAGGCTGCTACTCCCGTCC	1200
Db	1141	GGCGGCTGCGGACAGAAATGATTTGCTGCGTGGTGGCGGAGGCTGCTACTCCCGTCC	1200
Qy	1201	GTTGTCTCAGCCAAATGGCGGAGCGGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG	1260
Db	1201	GTTGTCTCAGCCAAATGGCGGAGCGGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG	1260

301 AGACCTACCAAGCTGGGCGCGCGCTAAACCGCGTGGCTCCGCCCATGACACCCCG 360
Db |||||
301 AGACCTACCAAGCTGGGCGCGCGCTAAACCGCGTGGCTCCGCCCATGACACCCCG 360
Qy |||||
361 CCAGTGCCTGATGCTGACCTCCGCGCGCGCATCTTGGCGCGGAGATTAACCTATCAACA 420
Db |||||
361 CCAGTGCCTGATGCTGACCTCCGCGCGCGCATCTTGGCGCGGAGATTAACCTATCAACA 420
Qy |||||
421 TCTCCCTTTACCTCTTCCGTGGCCACCGGACCTAACCTGCTTCTTTATGCCGCCCTCTT 480
Db |||||
421 TCTCCCTTTACCTCTTCCGTGGCCACCGGACCTAACCTGCTTCTTTATGCCGCCCTCTT 480
Qy |||||
481 AGTCGCTTTTACCCCTTCAGGACGCGCAACATACCATATATATGCGCCACGGAAGCTTCT 540
Db |||||
481 AGTCGCTTTTACCCCTTCAGGACGCGCAACATACCATATATATGCGCCACGGAAGCTTCT 540
Qy |||||
541 AATTATGCCAGTACCGGCTTGGCGGTGCGCAATCCGTTACCGCGCTGGTTCGCCAAT 600
Db |||||
541 AATTATGCCAGTACCGGCTTGGCGGTGCGCAATCCGTTACCGCGCTGGTTCGCCAAT 600
Qy |||||
601 GCTGTCCGCGGTTACGCCATCTCCATCTCATTTCTGSCCAGACACCAACCCCGGAG 660
Db |||||
601 GCTGTCCGCGGTTACGCCATCTCCATCTCATTTCTGSCCAGACACCAACCCCGGAG 660
Qy |||||
661 TCCGTTGATATGAATTCATTAACCTCGACGAGTGTCTGTTATTTAGTCAGGCCCGGCATA 720
Db |||||
661 TCCGTTGATATGAATTCATTAACCTCGACGAGTGTCTGTTATTTAGTCAGGCCCGGCATA 720
Qy |||||
721 GCCTGTGAGCTGTGATCCCAAGTAGCGGCTACACTATCGTAACCAAGGCTGGCGTCC 780
Db |||||
721 GCCTGTGAGCTGTGATCCCAAGTAGCGGCTACACTATCGTAACCAAGGCTGGCGTCC 780
Qy |||||
781 GTCGAGACTCTGGGGTGGCTGAGGAGGAGGCTACTCTGCTGTTGTTATGCTTTGCATA 840
Db |||||
781 GTCGAGACTCTGGGGTGGCTGAGGAGGAGGCTACTCTGCTGTTGTTATGCTTTGCATA 840
Qy |||||
841 CATGGCTCACTCGTAAATTCCTATATAACACCCCTATACCGGTGCCCTCGGGCTGTTG 900
Db |||||
841 CATGGCTCACTCGTAAATTCCTATATAACACCCCTATACCGGTGCCCTCGGGCTGTTG 900
Qy |||||
901 GACTTTGGCTTGAGCTTGAGTTTGGAACTTACCCCGGTAACACCAATACGCGGCTC 960
Db |||||
901 GACTTTGGCTTGAGCTTGAGTTTGGAACTTACCCCGGTAACACCAATACGCGGCTC 960
Qy |||||
961 TCCGCTTATTCAGCACTGCTGCCACCGCTTCTGTCGGGTGGGAGCGGACTGCCAG 1020
Db |||||
961 TCCGCTTATTCAGCACTGCTGCCACCGCTTCTGTCGGGTGGGAGCGGACTGCCAG 1020
Qy |||||
1021 CTCACCAACCGGCTGTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGGT 1080
Db |||||
1021 CTCACCAACCGGCTGTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGGT 1080
Qy |||||
1081 GTCGGTGAGATCGCGCGGAGTAGCCCTCACCTGTTTCAACCTTGTGACACTCTGTT 1140
Db |||||
1081 GTCGGTGAGATCGCGCGGAGTAGCCCTCACCTGTTTCAACCTTGTGACACTCTGTT 1140
Qy |||||
1141 GCGGCTTCCGACAGAAATGATTTGCTGCGCTGGTGGCCAGCTGTTCTCTCCGCTCC 1200
Db |||||
1141 GCGGCTTCCGACAGAAATGATTTGCTGCGCTGGTGGCCAGCTGTTCTCTCCGCTCC 1200
Qy |||||
1201 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db |||||
1201 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Qy |||||
1261 CAGGATAAGGTTATGCAATCCCGCATGACATGACCTCGGAGAACTCTCGTGGTTATT 1320
Db |||||
1261 CAGGATAAGGTTATGCAATCCCGCATGACATGACCTCGGAGAACTCTCGTGGTTATT 1320
Qy |||||
1321 CAGGATTATGATAACCAACATGACAGATCGCGGAGCTTCTCCAGCCCACTCGGC 1380
Db |||||
1321 CAGGATTATGATAACCAACATGACAGATCGCGGAGCTTCTCCAGCCCACTCGGC 1380
Qy |||||
1381 CCTTTCTGTCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCCAGTAT 1440

Db |||||
1381 CCTTTCTGTCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCCAGTAT 1440
Qy |||||
1441 GACCACTCACTATAGCTTCTTGCAGTGGCCAGATTTATGTTTCTGACTCTGTGACCTTG 1500
Db |||||
1441 GACCACTCACTATAGCTTCTTGCAGTGGCCAGATTTATGTTTCTGACTCTGTGACCTTG 1500
Qy |||||
1501 GTTAATGTTGACACCGGCGCGGAGCGGCTTGGCCGGTGGCTCGATTGGACCAAGGTCA 1560
Db |||||
1501 GTTAATGTTGACACCGGCGCGGAGCGGCTTGGCCGGTGGCTCGATTGGACCAAGGTCA 1560
Qy |||||
1561 CTTGACGGTTCGCGCCCTCTTCCACCACTCCAGCAGTACTCCAGAGACCTTCTTTGTCTGCGG 1620
Db |||||
1561 CTTGACGGTTCGCGCCCTCTTCCACCACTCCAGCAGTACTCCAGAGACCTTCTTTGTCTGCGG 1620
Qy |||||
1621 CTTCCGCGGTAAGCTCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCTTTATAAT 1680
Db |||||
1621 CTTCCGCGGTAAGCTCTCTTTCTGGGAGGAGGACCAACTAAAGCCGGGTACCTTTATAAT 1680
Qy |||||
1681 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATGCCCGCGGACCGGGTGGCT 1740
Db |||||
1681 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATGCCCGCGGACCGGGTGGCT 1740
Qy |||||
1741 ATTTCCACTTACCACTAGCTGCTGGGTGCTGGTCCGCTCTCCATTTCTGCGGTTCGCGTT 1800
Db |||||
1741 ATTTCCACTTACCACTAGCTGCTGGGTGCTGGTCCGCTCTCCATTTCTGCGGTTCGCGTT 1800
Qy |||||
1801 TTAGCCCCCACTGCTGCTAGCATTGCTTGGAGGATACCTTGGACTACCTTCCCGCGCC 1860
Db |||||
1801 TTAGCCCCCACTGCTGCTAGCATTGCTTGGAGGATACCTTGGACTACCTTCCCGCGCC 1860
Qy |||||
1861 CATACTTTTGTGATGATTTCTGCCAGAGTCCGCCCTTGAAGATGAAGTGGTAAAGTTCGGAGTTG 1920
Db |||||
1861 CATACTTTTGTGATGATTTCTGCCAGAGTCCGCCCTTGAAGATGAAGTGGTAAAGTTCGGAGTTG 1920
Qy |||||
1921 CAGTCTACTGCTGCTGAGCTTACGCGCTTGAAGATGAAGTGGTAAAGTTCGGAGTTG 1980
Db |||||
1921 CAGTCTACTGCTGCTGAGCTTACGCGCTTGAAGATGAAGTGGTAAAGTTCGGAGTTG 1980
Qy |||||
1981 TAG 1983
Db |||||
1981 TAG 1983

RESULT 4

US-07-876-941A-3
; Sequence 3, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876,941A
; FILING DATE: 01-MAY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.33
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2049 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 7
US-07-876-941A-3

Query Match 100.0%; Score 1983; DB 2; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
DB 1 ATGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
QY 61 CCGCGCGGTGAGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCG 120
DB 61 CCGCGCGGTGAGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCG 120
QY 121 TTCTGGGGTGACCGGGTGGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
DB 121 TTCTGGGGTGACCGGGTGGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
QY 181 CCCTTTCGCCCCCGATGTACCGGTGCGCGCGGGGCTGGACCTCGTGTTCGCCAACCCGCC 240
DB 181 CCCTTTCGCCCCCGATGTACCGGTGCGCGCGGGGCTGGACCTCGTGTTCGCCAACCCGCC 240
QY 241 CGACCACTCGGCTCGGCTTGGGTGACACAGGCCCGCGCGCGCGGTGCGGTGCGGTGCG 300
DB 241 CGACCACTCGGCTCGGCTTGGGTGACACAGGCCCGCGCGCGCGGTGCGGTGCGGTGCG 300
QY 301 AGACCTACCAAGCTGGCGCGCGCGCTAACCGCGGTGCGTCCGGGCCCATGACACCCCG 360
DB 301 AGACCTACCAAGCTGGCGCGCGCGCTAACCGCGGTGCGTCCGGGCCCATGACACCCCG 360
QY 361 CCAGTGGCTGATGTGCACTCCCGCGGCCCATCTTGGCGCGGCAGTATATCACTATCAACA 420
DB 361 CCAGTGGCTGATGTGCACTCCCGCGGCCCATCTTGGCGCGGCAGTATATCACTATCAACA 420
QY 421 TCTCCCTTACCTCTTCGTTGGCCACCGGCACTAACCTGGTTCTTTATGCGCGCCCTCT 480
DB 421 TCTCCCTTACCTCTTCGTTGGCCACCGGCACTAACCTGGTTCTTTATGCGCGCCCTCT 480
QY 481 AGTCGGCTTTTACCCCTTCAGGACGGCAACCAATACCCATATATATGGCCACGGAGCTTCT 540

DB 481 AGTCGGCTTTTACCCCTTCAGGACGGCAACCAATACCCATATATATGGCCACGGAGCTTCT 540
QY 541 AATTATGCCAGTACCGGGTTCGCCGTGCCCAAAATCCGTTACCGCGCTGGTCCCAAT 600
DB 541 AATTATGCCAGTACCGGGTTCGCCGTGCCCAAAATCCGTTACCGCGCTGGTCCCAAT 600
QY 601 GCTGTCGGCGGTTACGCCATCTCCATCTCATCTGGCCACAGACACCAACCCCGGAGC 660
DB 601 GCTGTCGGCGGTTACGCCATCTCCATCTCATCTGGCCACAGACACCAACCCCGGAGC 660
QY 661 TCCGTTGATATGAATTTCAATAACCTCGACGGATGTTGTTATTTAGTCCAGCCCGGCATA 720
DB 661 TCCGTTGATATGAATTTCAATAACCTCGACGGATGTTGTTATTTAGTCCAGCCCGGCATA 720
QY 721 GCTCTGAGCTTGTATCCCAAGTGAAGCGCTACCTATCGTAACCAAGCGTGGCGCTCC 780
DB 721 GCTCTGAGCTTGTATCCCAAGTGAAGCGCTACCTATCGTAACCAAGCGTGGCGCTCC 780
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGGGTACCTCTGGTCTTGTATGCTTTGCATA 840
DB 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGGGTACCTCTGGTCTTGTATGCTTTGCATA 840
QY 841 CATGGCTCACTCGTAAATTCCTATATACTAATAACCCCTATACCGGTGCCCTCGGGCTGTTG 900
DB 841 CATGGCTCACTCGTAAATTCCTATATACTAATAACCCCTATACCGGTGCCCTCGGGCTGTTG 900
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATACCGGGTCC 960
DB 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATACCGGGTCC 960
QY 961 TCCCGTTATTTCAGCACTGCTCGCCACCGCTTCGTGCGGTGCGGACGGAGCTGCGGAG 1020
DB 961 TCCCGTTATTTCAGCACTGCTCGCCACCGCTTCGTGCGGTGCGGACGGAGCTGCGGAG 1020
QY 1021 CTCAACCAACCGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
DB 1021 CTCAACCAACCGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
QY 1081 GTCGGTGAGATCGCGCGGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
DB 1081 GTCGGTGAGATCGCGCGGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
QY 1141 GCGCGGCTGCGGACAGAAATGATTTGCTGCGGTGCGGCGAGCTGTTCTACTCCGTCCT 1200
DB 1141 GCGCGGCTGCGGACAGAAATGATTTGCTGCGGTGCGGCGAGCTGTTCTACTCCGTCCT 1200
QY 1201 GTTGTCTCAGCCAAATGGCGAGCGCACTGTTAAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
DB 1201 GTTGTCTCAGCCAAATGGCGAGCGCACTGTTAAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
QY 1261 CAGGATAAGGGTATTCGAATCCCGCATGACATTGACCTCGGAGAACTCTCGTGTGTTATT 1320
DB 1261 CAGGATAAGGGTATTCGAATCCCGCATGACATTGACCTCGGAGAACTCTCGTGTGTTATT 1320
QY 1321 CAGGATTATGATAACCAACATGAACAGATCGGCGGACGCTTCTCCAGCCCAATCGCGC 1380
DB 1321 CAGGATTATGATAACCAACATGAACAGATCGGCGGACGCTTCTCCAGCCCAATCGCGC 1380
QY 1381 CTTTCTCTGTCTTCAGCTAATGATGCTGTTGGCTCTCTCTCAGCGGTGCGGAGTAT 1440
DB 1381 CTTTCTCTGTCTTCAGCTAATGATGCTGTTGGCTCTCTCTCAGCGGTGCGGAGTAT 1440
QY 1441 GACCACTGCTTATGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
DB 1441 GACCACTGCTTATGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
QY 1501 GTTAATGTTGCAACCGCGCGAGCGGTGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1560
DB 1501 GTTAATGTTGCAACCGCGCGAGCGGTGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1560
QY 1561 CTTGACCGGTGCGCGCGCTCTCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCGCG 1620

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0093.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2049 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 7
US-07-870-985A-3

Query Match 100.0%; Score 1983; DB 4; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATCGCGCCTCGGCTATTTGTTGCTGCTCTCATGTTTGTCTATGCTATGCTGCGCGGCCA	60
DB	1	ATCGCGCCTCGGCTATTTGTTGCTGCTCTCATGTTTGTCTATGCTATGCTGCGCGGCCA	60
QY	61	CCGCCCGGTCAGCGCTGCGCGCGCGTCTGCGCGCGCGCGAGCGCGGTTCCGCGGTTGT	120
DB	61	CCGCCCGGTCAGCGCTGCGCGCGCGTCTGCGCGCGCGCGAGCGCGGTTCCGCGGTTGT	120
QY	121	TTCTGGGGTGACCGGGTTGATTTCTCAGCCTTCGCAATCCCTATATTCATCCAAACCAAC	180
DB	121	TTCTGGGGTGACCGGGTTGATTTCTCAGCCTTCGCAATCCCTATATTCATCCAAACCAAC	180
QY	181	CCCTTCGCGCCCGATGTACCGCTGCGCGCGCGGCTGGAATCTGTTTCGCCAACCCGCC	240
DB	181	CCCTTCGCGCCCGATGTACCGCTGCGCGCGCGGCTGGAATCTGTTTCGCCAACCCGCC	240
QY	241	CGACCACTCGGCTCGGCTTGGCGTGAACAGCCAGCGCGCGCGGTTGCTTCACTGCTGCT	300
DB	241	CGACCACTCGGCTCGGCTTGGCGTGAACAGCCAGCGCGCGCGGTTGCTTCACTGCTGCT	300
QY	301	AGACCTACCACTGCGCGCGCGCGCTAACCGCGTCTCGCGCGCGCGCGCGCGCGCGCG	360
DB	301	AGACCTACCACTGCGCGCGCGCGCGCTAACCGCGTCTCGCGCGCGCGCGCGCGCGCG	360
QY	361	CCAGTGCTGATGTGCACTCCCGCGCGCGCGCTTTCGCGCGCGCGAGTAACTTCAACA	420
DB	361	CCAGTGCTGATGTGCACTCCCGCGCGCGCGCTTTCGCGCGCGCGAGTAACTTCAACA	420
QY	421	TCTCCCTTACCTTCTCCGTTGGCGACCGCGGCTTCTTTATGCGCGCGCTCTT	480
DB	421	TCTCCCTTACCTTCTCCGTTGGCGACCGCGGCTTCTTTATGCGCGCGCTCTT	480

481 AGTCCGCTTTTACCCCTTCAGGACGSCACCAATACCCATATATATGCGCACCGAAGCTTCT 540
481 AGTCCGCTTTTACCCCTTCAGGACGSCACCAATACCCATATATATGCGCACCGAAGCTTCT 540
541 AATTATGCCAGTACCGGGTTGCCGTGCCCAATCCGTTACCGCGCGCTGGTCCCAAT 600
541 AATTATGCCAGTACCGGGTTGCCGTGCCCAATCCGTTACCGCGCGCTGGTCCCAAT 600
601 GCTGTCGGGGTTAGCCATCTCCATCTCAATTCCTGCGCACAGACACCAACCCCGAGC 660
601 GCTGTCGGGGTTAGCCATCTCCATCTCAATTCCTGCGCACAGACACCAACCCCGAGC 660
661 TCCGTTGATATGAATTCATTAACCTCGACGGATGTTGCTATTTTAGTCCAGCCCGCAT 720
661 TCCGTTGATATGAATTCATTAACCTCGACGGATGTTGCTATTTTAGTCCAGCCCGCAT 720
721 GCCTCTGAGCTTTGATCCCAAGTGAGCGCCTACACTATCGTAAACCAAGCTGGCGCTCC 780
721 GCCTCTGAGCTTTGATCCCAAGTGAGCGCCTACACTATCGTAAACCAAGCTGGCGCTCC 780
781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTGTTGTTATGCTTTCATA 840
781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTGTTGTTATGCTTTCATA 840
841 CATGGCTCACTCGTAAATTCCTATATAACCCCTATACCGGTGCGCTCGGGCTGTTG 900
841 CATGGCTCACTCGTAAATTCCTATATAACCCCTATACCGGTGCGCTCGGGCTGTTG 900
901 GACTTTGCCCTTGAGCTTCAGTTTGCACCTTACCCCGGTAAACCAATACCGCGGTC 960
901 GACTTTGCCCTTGAGCTTCAGTTTGCACCTTACCCCGGTAAACCAATACCGCGGTC 960
961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGCGGGTGGCGAGCGGAGTCCGCGAG 1020
961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGCGGGTGGCGAGCGGAGTCCGCGAG 1020
1021 CTCACCAACCGCTGCTACCCGCTTTATGAAGAGACTCTATTTTACTAGTAAATGGT 1080
1021 CTCACCAACCGCTGCTACCCGCTTTATGAAGAGACTCTATTTTACTAGTAAATGGT 1080
1081 GTCGGTGAATCGGCGCGGATAGCCCTCACCTGTTTCAACCTGCTGACACTCTGCTT 1140
1081 GTCGGTGAATCGGCGCGGATAGCCCTCACCTGTTTCAACCTGCTGACACTCTGCTT 1140
1141 GCGCGCTCGCGACAAATTTGATTTCTGCGGCTGGTGGCGAGCTGTTACTCCCGTCCC 1200
1141 GCGCGCTCGCGACAAATTTGATTTCTGCGGCTGGTGGCGAGCTGTTACTCCCGTCCC 1200
1201 GTTGTCTCAGCCAAATGGCGAGCGGACTGTTTAAAGTTGTATATATCTGTAGAGAATGCTCAG 1260
1201 GTTGTCTCAGCCAAATGGCGAGCGGACTGTTTAAAGTTGTATATATCTGTAGAGAATGCTCAG 1260
1261 CAGGATAAGGGTATTGCAATTCGCGATGACATTTGAGTCTGGAGAAATCTGTTGGTTATT 1320
1261 CAGGATAAGGGTATTGCAATTCGCGATGACATTTGAGTCTGGAGAAATCTGTTGGTTATT 1320
1321 CAGGATTATGATAACCAACATGAAGATCGCGCGAGCGCTTCTCAGCGCCCATCGCGC 1380
1321 CAGGATTATGATAACCAACATGAAGATCGCGCGAGCGCTTCTCAGCGCCCATCGCGC 1380
1381 CCTTTCTGCTCTCGAGCTAATGATGTTGGCTCTCTCTCAGCGCTGCGGAGTAT 1440
1381 CCTTTCTGCTCTCGAGCTAATGATGTTGGCTCTCTCTCAGCGCTGCGGAGTAT 1440
1441 GACAGTCACTTATGGCTTTCAGTGGCGCGAGTTTATGTTTCTGACTGTGACCTTG 1500
1441 GACAGTCACTTATGGCTTTCAGTGGCGCGAGTTTATGTTTCTGACTGTGACCTTG 1500
1501 GTTAAATGTCGACCGCGCGAGCGGCTTGGCGGCTCGCTCGATTCGATTCGACCAAGGTCA 1560
1501 GTTAAATGTCGACCGCGCGAGCGGCTTGGCGGCTCGCTCGATTCGATTCGACCAAGGTCA 1560

1561 CTTGACGGTCCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTCTTTGTCTGCGG 1620
Db CTTGACGGTCCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTCTTTGTCTGCGG 1620
1621 CTCGCGGTAAGCTCTCTTTCTGGAGGAGGACACACTAAAGCCGGTACCCCTTATAAT 1680
Db CTCGCGGTAAGCTCTCTTTCTGGAGGAGGACACACTAAAGCCGGTACCCCTTATAAT 1680
1681 TATAACACCACTCTAGCAGCAACACTGCTTGTGAGAAATCCCGCCGACCGGGTCGCT 1740
Db TATAACACCACTCTAGCAGCAACACTGCTTGTGAGAAATCCCGCCGACCGGGTCGCT 1740
1741 ATTTCACCTTACACCACTAGCCTGGGTGCTGCTCCCAATTTCTGCGGTTGCGGTT 1800
Db ATTTCACCTTACACCACTAGCCTGGGTGCTGCTCCCAATTTCTGCGGTTGCGGTT 1800
1801 TTAGCCCCCACTCTGGCTAGCAATTTGCTTGGAGATACCTTGACCTGCCCCGGCC 1860
Db TTAGCCCCCACTCTGGCTAGCAATTTGCTTGGAGATACCTTGACCTGCCCCGGCC 1860
1861 CATACTTTTGTGATGATTTCTGCCAGAGTGGCGCCCTTGGCCTTCAGGCTGCGCTTTC 1920
Db CATACTTTTGTGATGATTTCTGCCAGAGTGGCGCCCTTGGCCTTCAGGCTGCGCTTTC 1920
1921 CAGTCTACTGCTGAGCTTCAGCCCTTAAAGATGAAGTGGTAAACTCGGGAGTTG 1980
Db CAGTCTACTGCTGAGCTTCAGCCCTTAAAGATGAAGTGGTAAACTCGGGAGTTG 1980
1981 TAG 1983
Db TAG 1983

RESULT 8

PCT-US95-13703-1
; Sequence 1, Application PC/TUS9513703
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
; TITLE OF INVENTION: US THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13703
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:

; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
; INDIVIDUAL ISOLATE: ORF-2
PCT-US95-13703-1

Query Match 100.0%; Score 1983; DB 5; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTGTCTATGTCGCCGCGCA	60
Db	1	ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTGTCTATGTCGCCGCGCA	60
QY	61	CGCCCGGTCAGCGCTCGGCTGCGGCGGCGGCGGCGGTCGCGGCGTGT	120
Db	61	CGCCCGGTCAGCGCTCGGCTGCGGCGGCGGCGGCGGTCGCGGCGTGT	120
QY	121	TTCTGGGGTGACCGGGTTCATTTCTCAGCCCTTTCGCAATCCCTTATATTTCATCCAAAC	180
Db	121	TTCTGGGGTGACCGGGTTCATTTCTCAGCCCTTTCGCAATCCCTTATATTTCATCCAAAC	180
QY	181	CCCTTCGCCCGGATGTCACCGCTGCGGCGGCGGCTGGACCTCGTTCGCAACCCGCC	240
Db	181	CCCTTCGCCCGGATGTCACCGCTGCGGCGGCGGCTGGACCTCGTTCGCAACCCGCC	240
QY	241	CGACCACTCGGCTCGGCTTGGGCTGACACAGGCGGCGGCGGCTGCTGCTGCTGCTGCT	300
Db	241	CGACCACTCGGCTCGGCTTGGGCTGACACAGGCGGCGGCGGCTGCTGCTGCTGCTGCT	300
QY	301	AGACCTTACCACAGCTGGGCGCGGCTTAAACCGGCTGCTCCGCGGCTGCTGACACCCCG	360
Db	301	AGACCTTACCACAGCTGGGCGCGGCTTAAACCGGCTGCTCCGCGGCTGCTGACACCCCG	360
QY	361	CCAGTCTGCTGATGTCGACTCCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	420
Db	361	CCAGTCTGCTGATGTCGACTCCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	420
QY	421	TCCTCCCTTACCTCTTCGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	480
Db	421	TCCTCCCTTACCTCTTCGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	480
QY	481	AGTCCGCTTTTACCCCTTTCAGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	540
Db	481	AGTCCGCTTTTACCCCTTTCAGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	540
QY	541	AATTATGCCAGTACCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG	600
Db	541	AATTATGCCAGTACCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG	600
QY	601	GCTGCGGCGGTTACCGCTTTCAGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	660
Db	601	GCTGCGGCGGTTACCGCTTTCAGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	660
QY	661	TCCGTTGATGATGAATTCATTAACCTTCAGGAGGCTGCTGCTGCTGCTGCTGCTGCT	720
Db	661	TCCGTTGATGATGAATTCATTAACCTTCAGGAGGCTGCTGCTGCTGCTGCTGCTGCT	720
QY	721	GCCTCTGAGCTTGTGATCCCAAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	780
Db	721	GCCTCTGAGCTTGTGATCCCAAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	780
QY	781	GTCGAGACCTCTGGGCTGCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	840
Db	781	GTCGAGACCTCTGGGCTGCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	840
QY	841	CATGGCTCACTCGTAAATTCCTTACTAATACCCCTTACCTGCTGCTGCTGCTGCTGCT	900
Db	841	CATGGCTCACTCGTAAATTCCTTACTAATACCCCTTACCTGCTGCTGCTGCTGCTGCT	900
QY	901	GACTTTCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960
Db	901	GACTTTCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960
QY	961	TCCGTTTATCCAGACTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1020

NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,054
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 870,985
FILING DATE: 20-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.38
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2094 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: BURMA SEQUENCE, FIGURE 7

US-08-484-054-1
Query Match 100.0%; Score 1983; DB 1; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCGGCTATTTGTTGCTGCTCATGTTTTTGGCTATGCTGCGCGGCCA 60
Db |
46 ATGCGCCCTCGGCTATTTGTTGCTGCTCATGTTTTTGGCTATGCTGCGCGGCCA 105
QY 61 CCGCGCGCTCAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120
Db |
106 CCGCGCGCTCAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGT 165
QY 121 TTCTGCGGGTACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 180
Db |
166 TTCTGCGGGTACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 225

181 CCCTTCGCCCCCGATGTACCGCTGCGCGCGGGCTGGACCTCGTTGTCGCCAACCGCC 240
|
226 CCCTTCGCCCCCGATGTACCGCTGCGCGCGGGCTGGACCTCGTTGTCGCCAACCGCC 285
QY 241 CGACCACTCGGCTCGGCTTGGGCTGACAGGGCCAGGCGCCGCGCTGCGCTCACGTCG 300
Db |
286 CGACCACTCGGCTCGGCTTGGGCTGACAGGGCCAGGCGCCGCGCTGCGCTCACGTCG 345
QY 301 AGACCTACCACTGCGGCGCGCTAACCGGCTCGCTCCGCGCCCATGACACCCCG 360
Db |
346 AGACCTACCACTGCGGCGCGCTAACCGGCTCGCTCCGCGCCCATGACACCCCG 405
QY 361 CCAGTGCCTGATGTGACTCCGCGCGCGCATCTTTCGCGCGGCGAGTAACTATCAACA 420
Db |
406 CCAGTGCCTGATGTGACTCCGCGCGCGCATCTTTCGCGCGGCGAGTAACTATCAACA 465
QY 421 TCTCCCTTACCTTTCCTGCGCGCGCATCTTTCGCGCGGCTTTCGCGCGGCTCTT 480
Db |
466 TCTCCCTTACCTTTCCTGCGCGCGCATCTTTCGCGCGGCTTTCGCGCGGCTCTT 525
QY 481 AGTCCGCTTTTACCCCTTCAGGACGCGCAATACCATATATATGCGCACGGAAGTCT 540
Db |
526 AGTCCGCTTTTACCCCTTCAGGACGCGCAATACCATATATATGCGCACGGAAGTCT 585
QY 541 AATTATGCCAGTACCGGGTTGCGCGTGCACAATCCGTTACCGCGCGTGTCCCAAT 600
Db |
586 AATTATGCCAGTACCGGGTTGCGCGTGCACAATCCGTTACCGCGCGTGTCCCAAT 645
QY 601 GCTGCGGCGGTACGCGATCTCCATCTATTCTGCGCACAGACACACACCGCGCGAG 660
Db |
646 GCTGCGGCGGTACGCGATCTCCATCTATTCTGCGCACAGACACACACCGCGCGAG 705
QY 661 TCGGTTGATGAAATTCATTAACCTCGACGGATGTTGCTATTATTAGTCCAGCCCGCAT 720
Db |
706 TCGGTTGATGAAATTCATTAACCTCGACGGATGTTGCTATTATTAGTCCAGCCCGCAT 765
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATGTAACCAAGCTGGCGTCC 780
Db |
766 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATGTAACCAAGCTGGCGTCC 825
QY 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
Db |
826 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 885
QY 841 CATGGCTCACTGTAATTCCTATATAACACCTATACACCTATACCGGCTGGCGTGTG 900
Db |
886 CATGGCTCACTGTAATTCCTATATAACACCTATACCGGCTGGCGTGGTGTG 945
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATACCGGGTC 960
Db |
946 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATACCGGGTC 1005
QY 961 TCCGTTATTCAGCACTGCTGCGCACCGGCTTCGTCGGGTGCGGACGAGTCCGAG 1020
Db |
1006 TCCGTTATTCAGCACTGCTGCGCACCGGCTTCGTCGGGTGCGGACGAGTCCGAG 1065
QY 1021 CTCACCAACACGCTGCTACCGCTTTATGAAGAGCTCTATTATTACTAGTAAATGGT 1080
Db |
1066 CTCACCAACACGCTGCTACCGCTTTATGAAGAGCTCTATTATTACTAGTAAATGGT 1125
QY 1081 GTCGCTGAGATCGCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTGTGCT 1140
Db |
1126 GTCGCTGAGATCGCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTGTGCT 1185
QY 1141 GCGCGCTGCGGACAGAAATGATTTGCTGCGGTGGTGGCGAGCTTCTACTCCGCTCC 1200
Db |
1186 GCGCGCTGCGGACAGAAATGATTTGCTGCGGTGGTGGCGAGCTTCTACTCCGCTCC 1245
QY 1201 GTTGTCTCAGCAATGCGGACGAGCTGTTAAGTTGTATACATCTCTAGAGAATGCTCAG 1260
Db |
1246 GTTGTCTCAGCAATGCGGACGAGCTGTTAAGTTGTATACATCTCTAGAGAATGCTCAG 1305
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAACTCTCGTGGGTATT 1320

Db 1306 CAGGATAAGGATATTGCAATCCCGCATGACATTGACCTCGGAGATCTCGTGTGTTAT 1365
Qy 1321 CAGGATTATGATAACCAACATGAACAAGATCGCGCGAGCTTCTCCAGCCCATCGCGC 1380
Db 1366 CAGGATTATGATAACCAACATGAACAAGATCGCGCGAGCTTCTCCAGCCCATCGCGC 1425
Qy 1381 CTTTCTCTGTCCTTTCGAGCTAATGATGCTTTTGGCTCTCTCCTCAGCGCTGCCAGTAT 1440
Db 1426 CTTTCTCTGTCCTTTCGAGCTAATGATGCTTTTGGCTCTCTCCTCAGCGCTGCCAGTAT 1485
Qy 1441 GACCACTCCACTTATGCTCTTCGACTGCGCCAGCTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 1486 GACCACTCCACTTATGCTCTTCGACTGCGCCAGCTTTATGTTTCTGACTCTGTGACCTTG 1545
Qy 1501 GTTATGTTGGACCGCGCGGAGCGGCTGTCGCGCGGCTGCTGATGAGCAAGGTACA 1560
Db 1546 GTTATGTTGGACCGCGCGGAGCGGCTGTCGCGCGGCTGCTGATGAGCAAGGTACA 1605
Qy 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGG 1620
Db 1606 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGG 1665
Qy 1621 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGGAGCAACAATAAGCGGGTACCTTTATAT 1680
Db 1666 CTCGCGGTAAAGCTCTCTTTCTGGAGGAGGAGCAACAATAAGCGGGTACCTTTATAT 1725
Qy 1681 TATAACACACTGTAGGACCAACTGCTTGTGAGATGCGCGGGGACCGGTGCT 1740
Db 1726 TATAACACACTGTAGGACCAACTGCTTGTGAGATGCGCGGGGACCGGTGCT 1785
Qy 1741 ATTTCACCTACACACTAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1786 ATTTCACCTACACACTAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1845
Qy 1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTGGAGATACCTTGGACTACCTGCGCGCGC 1860
Db 1846 TTAGCCCCCACTCTGCGCTAGCAATGCTTGGAGATACCTTGGACTACCTGCGCGCGC 1905
Qy 1861 CATACCTTTGATGATTTCTGCGCAGAGTGCCTGCGCGCGCTTGGCTTTCAGGGTTC 1920
Db 1906 CATACCTTTGATGATTTCTGCGCAGAGTGCCTGCGCGCGCTTGGCTTTCAGGGTTC 1965
Qy 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGTAAACTCGGAGTTG 1980
Db 1966 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGTAAACTCGGAGTTG 2025
Qy 1981 TAG 1983
Db 2026 TAG 2028

RESULT 11
US-07-876-941A-1
; Sequence 1, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876.941A
; FILING DATE: 01-MAY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-00993.33
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2094 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BURMA SEQUENCE, FIGURE 7
US-07-876-941A-1

Query Match 100.0%; Score 1983; DB 2; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCCCTCGGCTTATTTTGTGCTGCTCTCATGTTTTTGGCTATGCTGCGCGGCA 60
Db 46 ATGGCGCCCTCGGCTTATTTTGTGCTGCTCTCATGTTTTTGGCTATGCTGCGCGGCA 105
Qy 61 CGCGCCGCTCAGCCGCTCTGGCGCGCGCTGCTGGCGCGCGAGCGGTTCCGGCGGTGGT 120
Db 106 CGCGCCGCTCAGCCGCTCTGGCGCGCGCTGCTGGCGCGCGAGCGGTTCCGGCGGTGGT 165
Qy 121 TTCTGGGTGACCGGTTGATTTCTCAGCCCTTCCGAATCCCTATATTTATCCAAACCAAC 180
Db 166 TTCTGGGTGACCGGTTGATTTCTCAGCCCTTCCGAATCCCTATATTTATCCAAACCAAC 225
Qy 181 CCCTTTCGCCCCCGATGTCCAGCTGCGCGCGGCTGAGCTGCTGTTTGGCAACCCGCC 240
Db 226 CCCTTTCGCCCCCGATGTCCAGCTGCGCGCGGCTGAGCTGCTGTTTGGCAACCCGCC 285
Qy 241 CGACCACTCGGCTCGGCTTGGCGTGAACAGCGCCCGCGCGCTTGGCTTCCCTCAGCTGCT 300
Db 286 CGACCACTCGGCTCGGCTTGGCGTGAACAGCGCCCGCGCGCTTGGCTTCCCTCAGCTGCT 345
Qy 301 AGACCTACCAAGCTGGGGCGCGCGCTAAACCGGCTGCTCCGGCCCATGACACCCCG 360
Db 346 AGACCTACCAAGCTGGGGCGCGCGCTAAACCGGCTGCTCCGGCCCATGACACCCCG 405
Qy 361 CCAGTGCCTGATGTCGACTCCCGCGGCGCACTTTTGGCGCGGAGTATAACCTATCAACA 420

Db 406 CCAGTGCCTGATGTGACCTCCGCGGCGCACTCTTGGCGGCGAGTATACCTATCAACA 465
QY 421 TCTCCCTTAACTCTTTCCTGTCGACACCGGCACTAACTCTGGTCTCTTTATGCGCGCCCTCTT 480
Db 466 TCTCCCTTAACTCTTTCCTGTCGACACCGGCACTAACTCTGGTCTCTTTATGCGCGCCCTCTT 525
QY 481 AGTCCGCTTTTACCCCTTCAGACGCGCACCAATACCATATATATGACGCGAGCTTCT 540
Db 526 AGTCCGCTTTTACCCCTTCAGACGCGCACCAATACCATATATATGACGCGAGCTTCT 585
QY 541 AATTATGCCAGTACCGGCTTGCCTGTCACAACTCCGTTACCGCCCGCTGGTCCCAAT 600
Db 586 AATTATGCCAGTACCGGCTTGCCTGTCACAACTCCGTTACCGCCCGCTGGTCCCAAT 645
QY 601 GCTGTCGCGGTTACGCCATCTCATCTCTGTCGCGCACAGACCAACACCGCGAGC 660
Db 646 GCTGTCGCGGTTACGCCATCTCATCTCTGTCGCGCACAGACCAACACCGCGAGC 705
QY 661 TCCGTTGATATGAATTCATTAACCTCGAGGATGTTGCTATTTTATGTCAGCGCGGCATA 720
Db 706 TCCGTTGATATGAATTCATTAACCTCGAGGATGTTGCTATTTTATGTCAGCGCGGCATA 765
QY 721 GCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGCTGCGCTCC 780
Db 766 GCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGCTGCGCTCC 825
QY 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGCTACCTCTGGCTTGTATGCTTTCGATA 840
Db 826 GTCGAGACCTCTGGGCTGCTGAGGAGGAGCTACCTCTGGCTTGTATGCTTTCGATA 885
QY 841 CATGGCTCATCTGTAATTCCTATCTATATACACCTATACCGGCTCGCGGCTGTTG 900
Db 886 CATGGCTCATCTGTAATTCCTATCTATATACACCTATACCGGCTCGCGGCTGTTG 945
QY 901 GACTTTGCGCTTGGAGCTTGGATTTGCAACTTACCCCGGTAACCAATACCGCGGTC 960
Db 946 GACTTTGCGCTTGGAGCTTGGATTTGCAACTTACCCCGGTAACCAATACCGCGGTC 1005
QY 961 TCCCGTTATTCAGACCTGCTCGCCACCGCTTCTGTCGCGGTCGCGAGCTGCGGAG 1020
Db 1006 TCCCGTTATTCAGACCTGCTCGCCACCGCTTCTGTCGCGGTCGCGAGCTGCGGAG 1065
QY 1021 CTCACACACAGCTGCTACCGCTTATGAAGGACTCTATTTTACTAGTACTAATGTT 1080
Db 1066 CTCACACACAGCTGCTACCGCTTATGAAGGACTCTATTTTACTAGTACTAATGTT 1125
QY 1081 GTCGCTGAGATCGCGCGGGATAGCCCTCACCTGTTCACCTTGTGACACTCTGCTT 1140
Db 1126 GTCGCTGAGATCGCGCGGGATAGCCCTCACCTGTTCACCTTGTGACACTCTGCTT 1185
QY 1141 GCGCGCTGCGGACAGATGATTTGTCGCGCTGGTGGCCAGCTGTTCTACTCCGCTCC 1200
Db 1186 GCGCGCTGCGGACAGATGATTTGTCGCGCTGGTGGCCAGCTGTTCTACTCCGCTCC 1245
QY 1201 GTTGTCTCAGCAGGCGAGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db 1246 GTTGTCTCAGCAGGCGAGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1305
QY 1261 CAGGATAAGGGTATTCGAATCCCGCATGACATTCAGCTCGGAAATCTCGTGGTTAT 1320
Db 1306 CAGGATAAGGGTATTCGAATCCCGCATGACATTCAGCTCGGAAATCTCGTGGTTAT 1365
QY 1321 CAGGATATGATTAACCAAGTACAGATGAGCGGCGAGCTTCTCAGCGCCCATCGGC 1380
Db 1366 CAGGATATGATTAACCAAGTACAGATGAGCGGCGAGCTTCTCAGCGCCCATCGGC 1425
QY 1381 CCTTCTCTCTCTCGAGCTTAATGATGCTTGGCTCTCTCTCAGCGCTGCGGAGTAT 1440
Db 1426 CCTTCTCTCTCTCGAGCTTAATGATGCTTGGCTCTCTCTCAGCGCTGCGGAGTAT 1485
QY 1441 GACCACTCTTATGCTCTTCCAGCTGGCCGAGTTATGTTCTGACTCTGAGCTTGG 1500

Db 1486 GACCACTCTTATGCTCTTCCAGCTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1545
QY 1501 GTTATATGTCGACCGGCGCGAGCGCTTGGCCCGCTCGATTTGGACCAAGGTCACA 1560
Db 1546 GTTATATGTCGACCGGCGCGAGCGCTTGGCCCGCTCGATTTGGACCAAGGTCACA 1605
QY 1561 CTTGAGCGTGGCGCCCTCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTCTGCGG 1620
Db 1606 CTTGAGCGTGGCGCCCTCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTCTGCGG 1665
QY 1621 CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGAGGAGCAAACTAAAGCCGGGTACCTTTATAAT 1680
Db 1666 CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGAGGAGCAAACTAAAGCCGGGTACCTTTATAAT 1725
QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCCCGGAGCAACCGGTCGCT 1740
Db 1726 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCCCGGAGCAACCGGTCGCT 1785
QY 1741 ATTTCACCTTACACCACTAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1786 ATTTCACCTTACACCACTAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1845
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGGAGTACCTTGGACTACCTGCGCGGCC 1860
Db 1846 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGGAGTACCTTGGACTACCTGCGCGGCC 1905
QY 1861 CATACTTTTGTGATTTCTGCCCCAGAGTCCGCCCCCTTTGGCTTCAGGGCTGCGCTTTC 1920
Db 1906 CATACTTTTGTGATTTCTGCCCCAGAGTCCGCCCCCTTTGGCTTCAGGGCTGCGCTTTC 1965
QY 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGATGAAGTGGGTAAAACTCGGGAGTTG 1980
Db 1966 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGATGAAGTGGGTAAAACTCGGGAGTTG 2025
QY 1981 TAG 1983
Db 2026 TAG 2028

RESULT 12

US-07-870-985A-1
; Sequence 1, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Two, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888


```
/ FILING DATE: 05-APRIL-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 420,921
/ FILING DATE: 13-OCTOBER-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 367,486
/ FILING DATE: 16-JUNE-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 336,672
/ FILING DATE: 11-APRIL-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 208,997
/ FILING DATE: 17-JUNE-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0093.30
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2094 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA to mRNA
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: BURMA SEQUENCE, FIGURE 7
US-07-870-985A-1

Query Match 100.0%; Score 1983; DB 4; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTTCCTATGTTTTCCTATGCTGCGCCGCGCA 60
DB 1 ATGGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTTCCTATGTTTTCCTATGCTGCGCCGCGCA 105
QY 61 CGCCCGGTGAGCGCTCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 1 CGCCCGGTGAGCGCTCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 165
QY 121 TTCTGGGTGACCGGTGATTTCTCAGCCCTTCGCAATCCCTATATTTATCCACCAAC 180
DB 1 TTCTGGGTGACCGGTGATTTCTCAGCCCTTCGCAATCCCTATATTTATCCACCAAC 225
QY 181 CCCTTCGCCCGCGATGTACCGCTCGCGCGCGCGCTGACCTCGTGTTCGCCCAACCCGCC 240
DB 226 CCCTTCGCCCGCGATGTACCGCTCGCGCGCGCGCTGACCTCGTGTTCGCCCAACCCGCC 285
QY 241 CGACACTCGGCTCGCTTGGGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 286 CGACACTCGGCTCGCTTGGGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 345
QY 301 AGACCTACACAGCTGGGCGCGCGCTAACCGCGGTGCGTTCGGCGCGCGCGCGCGCGCGCGCG 360
DB 346 AGACCTACACAGCTGGGCGCGCGCTAACCGCGGTGCGTTCGGCGCGCGCGCGCGCGCGCGCG 405
QY 361 CCAGTGCCTGATGTGCACTCCCGCGCGCGCTTTCGCCCGCGCGATATAACCTATCAACA 420
DB 406 CCAGTGCCTGATGTGCACTCCCGCGCGCGCTTTCGCCCGCGCGATATAACCTATCAACA 465
QY 421 TCTCCCTTACCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 480
DB 466 TCTCCCTTACCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 525
QY 481 AGTCGGCTTTTACCCCTTCAGGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 526 AGTCGGCTTTTACCCCTTCAGGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 585
QY 541 AATTATGCCAGTACCGGGTTCGCGGTGCCAATCCCGTTCACCGCGCGCGCGCGCGCGCGCGCG 600
```

```
DB 586 AATTATGCCAGTACCGGGTTCGCGGTGCCAATCCCGTTCACCGCGCGCGTGTGCCAAT 645
QY 601 GCTGTGCGGGTTACGCATCTCCATCTCATTTCTGGCCACAGACCAACCAACCGCGCGCG 660
DB 646 GCTGTGCGGGTTACGCATCTCCATCTCATTTCTGGCCACAGACCAACCAACCGCGCGCG 705
QY 661 TCCGTTGATATGAATTAATTAACCTCGAGCGGATGTTGTTATTTTATGTCAGCGCGCGCAT 720
DB 706 TCCGTTGATATGAATTAATTAACCTCGAGCGGATGTTGTTATTTTATGTCAGCGCGCGCAT 765
QY 721 GCCTCTGAGCTTGTGATCCCAAGTACGCGCTTACACTATCGTAACCAAGGCTGGCGCTCC 780
DB 766 GCCTCTGAGCTTGTGATCCCAAGTACGCGCTTACACTATCGTAACCAAGGCTGGCGCTCC 825
QY 781 GTCGAGACTCTGCGGGTGGCTGAGGAGGAGGCTACCTCTGTCCTTTTATGTTTATGTCATA 840
DB 826 GTCGAGACTCTGCGGGTGGCTGAGGAGGAGGCTACCTCTGTCCTTTTATGTTTATGTCATA 885
QY 841 CATGGCTCACTCGTAAATTCCTATATACTAATACACCTATACCGGTGCGCTCGGGCTGTTG 900
DB 886 CATGGCTCACTCGTAAATTCCTATATACTAATACACCTATACCGGTGCGCTCGGGCTGTTG 945
QY 901 GACTTTGCGCTTGAAGTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGCTC 960
DB 946 GACTTTGCGCTTGAAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGCTC 1005
QY 961 TCCGTTATTCAGACTGTCTGCGCACCGCTTGTGCGGTGGGAGCGGAGTCCCGAG 1020
DB 1006 TCCGTTATTCAGACTGTCTGCGCACCGCTTGTGCGGTGGGAGCGGAGTCCCGAG 1065
QY 1021 CTCACCAACGCTGTACCGCTTATGAGGAGCTCTATTTTACTAGTACTAATGTT 1080
DB 1066 CTCACCAACGCTGTACCGCTTATGAGGAGCTCTATTTTACTAGTACTAATGTT 1125
QY 1081 GTCGCTGAGATCGCGCGGAGTACCTCTCACTCCCTGTTCAACCTTGTGACACTCTGCTT 1140
DB 1126 GTCGCTGAGATCGCGCGGAGTACCTCTCACTCCCTGTTCAACCTTGTGACACTCTGCTT 1185
QY 1141 GCGGCGCTGCGACAGAAATGATTTGTCGCTGCTGTCGCGAGCTGTTTACTCCTCGTCC 1200
DB 1186 GCGGCGCTGCGACAGAAATGATTTGTCGCTGCTGTCGCGAGCTGTTTACTCCTCGTCC 1245
QY 1201 GTTCTCTCAGCAATGCGCGCGAGCTGTTAAGTTGTATACATCTGTAGAGAGTCTCAG 1260
DB 1246 GTTGTCTCAGCAATGCGCGAGCTGTTAAGTTGTATACATCTGTAGAGAGTCTCAG 1305
QY 1261 CAGGATAAGGCTATTTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGTTAT 1320
DB 1306 CAGGATAAGGCTATTTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGTTAT 1365
QY 1321 CAGGATAAGGCTATTTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGTTAT 1380
DB 1366 CAGGATAAGGCTATTTGCAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGTTAT 1425
QY 1381 CCTTCTCTGCTTCGAGCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
DB 1426 CCTTCTCTGCTTCGAGCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1485
QY 1441 GACCAGTCCACTTATGCTCTTCCAGCTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTGT 1500
DB 1486 GACCAGTCCACTTATGCTCTTCCAGCTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTGT 1545
QY 1501 GTTAAATGTTGCGACGCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
DB 1546 GTTAAATGTTGCGACGCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1605
QY 1561 CTTGACCGTTCGCGCGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCTGCTGCT 1620
DB 1606 CTTGACCGTTCGCGCGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCTGCTGCT 1665
QY 1621 CTTGACCGTTCGCGCGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCTGCTGCT 1680
```

Db 1666 CTCGCGGTAGACTCTCTTTCTGGGAGGCGAGCAAACTAAAGCCGGTACCTTATAAT 1725
QY 1681 TATAACACACTGCTAGGACCAACTGCTTGTGAGAAATGCCGCGGACCGGGTCTGCT 1740
Db 1726 TATAACACCACTGCTAGGACCAACTGCTTGTGAGAAATGCCGCGGACCGGGTCTGCT 1785
QY 1741 ATTTCACCTTACACCACTAGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1786 ATTTCACCTTACACCACTAGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1845
QY 1801 TTAGCCCCCACTCTGCTAGCACTTGTGAGGATACCTTGAGACTACCTGCGCGGCC 1860
Db 1846 TTAGCCCCCACTCTGCTAGCACTTGTGAGGATACCTTGAGACTACCTGCGCGGCC 1905
QY 1861 CATACTTTTATGATTTCTGCCAGAGTGCAGGCTGCGGCTTTCAGGCTGCGGCTTTC 1920
Db 1906 CATACTTTTATGATTTCTGCCAGAGTGCAGGCTGCGGCTTTCAGGCTGCGGCTTTC 1965
QY 1921 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTTCAGGCTGCGGCTTTCAGGCTGCGGCTTTC 1980
Db 1966 CAGTCTACTGCTGCTGAGCTTTCAGCGCTTTCAGGCTGCGGCTTTCAGGCTGCGGCTTTC 2025
QY 1981 TAG 1983
Db 2026 TAG 2028

RESULT 13
US-08-478-507-6
; Sequence 6, Application US/08478507
; Patent No. 6120988
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarborough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6120988-A/No. 6120988-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672

; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0183.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HEV - Burma strain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..5106
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5147..7126
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5106..5474
; US-08-478-507-6

Query Match 100.0%; Score 1983; DB 3; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGGCTATTTTGTGCTCCTCATGTTTGTGCTATGCTGCGCGGCCA 60
Db 5147 ATGCGCCCTCGGCTATTTTGTGCTCCTCATGTTTGTGCTATGCTGCGCGGCCA 5206

QY 61 CCGCCCGGTGACCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120
Db 5207 CCGCCCGGTGACCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 5266

QY 121 TTCTGGGTGACCGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 180
Db 5267 TTCTGGGTGACCGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 5326

QY 181 CCTTTCGCGCCCGATGTCACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 5327 CCTTTCGCGCCCGATGTCACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5386

QY 241 CGACCACTCGGCTCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 300
Db 5387 CGACCACTCGGCTCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 5446

QY 301 AGACCTACACAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 5447 AGACCTACACAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5506

QY 361 CCAGTGCCTGATGTCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 420
Db 5507 CCAGTGCCTGATGTCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 5566

QY 421 TCTCCCTTACCTCTTCCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 480
Db 5567 TCTCCCTTACCTCTTCCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 5626

QY 481 AGTCCGCTTTTACCCCTTCAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 540
Db 5627 AGTCCGCTTTTACCCCTTCAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGT 5686

QY 541 AATTATGCCAGTACCGGGTGGCCGTCGACCAATCCGTTACCGCCGCGTGGTCCCAAT 600
Db 5687 AATTATGCCAGTACCGGGTGGCCGTCGACCAATCCGTTACCGCCGCGTGGTCCCAAT 5746
QY 601 GCTGTCCGGGTACCGCATCTCCATCTCAATCTTGTGGCCACAGACCAACACCCCGAAG 660
Db 5747 GCTGTCCGGGTACCGCATCTCCATCTCAATCTTGTGGCCACAGACCAACACCCCGAAG 5806
QY 661 TCCGTTGATATGAATCAATTAACCTCCAGCGATCTTGTAATTTAGTCCAGCCGGCATA 720
Db 5807 TCCGTTGATATGAATCAATTAACCTCCAGCGATCTTGTAATTTAGTCCAGCCGGCATA 5866
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGTGCGCTCC 780
Db 5867 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGTGCGCTCC 5926
QY 781 GTCGAGACCTCTGGGGTGGTGAGGAGAGGCTACCTCTGGTCTTGTATGTCTTTGATA 840
Db 5927 GTCGAGACCTCTGGGGTGGTGAGGAGAGGCTACCTCTGGTCTTGTATGTCTTTGATA 5986
QY 841 CATGGCTCACTCGTAATTCCTATCTATACTAATAACCCCTATACCGTGCCCTCGGCTGTG 900
Db 5987 CATGGCTCACTCGTAATTCCTATCTATACTAATAACCCCTATACCGTGCCCTCGGCTGTG 6046
QY 901 GACTTTGCCCTTGAGCTTGATTTGCAACCTTACCCCGGTAAACCAATACCGGGTC 960
Db 6047 GACTTTGCCCTTGAGCTTGATTTGCAACCTTACCCCGGTAAACCAATACCGGGTC 6106
QY 961 TCCGTTATTCAGACTGCTCGCCACCGCTTCTGTCGGTGCGGACGGAGCTGCCGAG 1020
Db 6107 TCCGTTATTCAGACTGCTCGCCACCGCTTCTGTCGGTGCGGACGGAGCTGCCGAG 6166
QY 1021 CTCACCAACCGCTGCTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAAGT 1080
Db 6167 CTCACCAACCGCTGCTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAAGT 6286
QY 1081 GTCGGTGAGATCGCCGCGGATAGCCCTCACCCCTGTTCAACCTTGCTGACACTCTGCT 1140
Db 6227 GTCGGTGAGATCGCCGCGGATAGCCCTCACCCCTGTTCAACCTTGCTGACACTCTGCT 6286
QY 1141 GCGGCTCGGACAGAAATGATTTGCTGGCTGGTGCGCAGCTGTTCTACTCCGCTCC 1200
Db 6287 GCGGCTCGGACAGAAATGATTTGCTGGCTGGTGCGCAGCTGTTCTACTCCGCTCC 6346
QY 1201 GTTCTCTCAGCAATGCGAGCCAGTGAAGTTGTATACATCTGTAGAGAAGTCTCAG 1260
Db 6347 GTTCTCTCAGCAATGCGAGCCAGTGAAGTTGTATACATCTGTAGAGAAGTCTCAG 6406
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT 1320
Db 6407 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT 6466
QY 1321 CAGGATTATGATACCAACATGAACAGATCGCGGAGCCCTTCTCAGCCCCCATCGGC 1380
Db 6467 CAGGATTATGATACCAACATGAACAGATCGCGGAGCCCTTCTCAGCCCCCATCGGC 6526
QY 1381 CTTTCTCTGCTTCTGAGCTTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT 1440
Db 6527 CTTTCTCTGCTTCTGAGCTTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGAGTAT 6586
QY 1441 GACCAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGATCTGTGACCTTG 1500
Db 6587 GACCAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCTGATCTGTGACCTTG 6646
QY 1501 GTTAATGTTGGACGGCGCGCAGCGCTTCCCGGTGCTCGATTTGAGCAACAGGTGACA 1560
Db 6647 GTTAATGTTGGACGGCGCGCAGCGCTTCCCGGTGCTCGATTTGAGCAACAGGTGACA 6706
QY 1561 CTTGACCGTCCGCCCTCTCCACCATCCAGAGTACTCGAGAGCTTCTTTGCTCCTCGC 1620
Db 6707 CTTGACCGTCCGCCCTCTCCACCATCCAGAGTACTCGAGAGCTTCTTTGCTCCTCGC 6766
QY 1621 CTCGCGGTAAAGCTCTCTTTCTGGAGGCGAGCAACCTAAAGCGGGTACCTTATAAT 1680

Db 6767 CTCGCGGTAAAGCTCTCTTTCTGGAGGCGAGCAACCTAAAGCGGGTACCTTATAAT 6826
QY 1581 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCCGCGGACCGGGTGGCT 1740
Db 6827 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCCGCGGACCGGGTGGCT 6886
QY 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTTGCCGTT 1800
Db 6887 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTTGCCGTT 6946
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTCTTGAGGATACCTTGGACTACCTGCCCGGCC 1860
Db 6947 TTAGCCCCCACTCTGCGCTAGCATTTCTTGAGGATACCTTGGACTACCTGCCCGGCC 7006
QY 1861 CATACTTTTATGATTTCTGCCAGAGTGGCGCCCTTGGCCCTTACGGGCTGGCTTTC 1920
Db 7007 CATACTTTTATGATTTCTGCCAGAGTGGCGCCCTTGGCCCTTACGGGCTGGCTTTC 7066
QY 1921 CAGTCTACTGCTGCTGAGCTTACAGCGCTTAAAGATGAAGTGGGTAACCTCGGAGTTG 1980
Db 7067 CAGTCTACTGCTGCTGAGCTTACAGCGCTTAAAGATGAAGTGGGTAACCTCGGAGTTG 7126
QY 1981 TAG 1983
Db 7127 TAG 7129

RESULT 14
US-09-128-275A-6
; Sequence 6, Application US/09128275A
; Patent No. 6229005
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6229005-A/No. 6229005-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,275A
; FILING DATE: 03-AUG-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HEV - Burma strain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28...5106
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5147..7126
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5106..5474
; US-09-128-275A-6

Query Match 100.0%; Score 1983; DB 3; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCGCCGCTATTTGTTGCTGCTCCTCATGTTTGTGCTATGCTGCGCCGCGCA 60
DB 5147 ATGCGCCCTCGCCGCTATTTGTTGCTGCTCCTCATGTTTGTGCTATGCTGCGCCGCGCA 5206

QY 61 CGCGCCCGGTGACCGGCTGCTGCGCGCGCGTCTGCGCGCGCGCGAGCGCGGTTCCGCGCGTGT 120
DB 5207 CGCGCCCGGTGACCGGCTGCTGCGCGCGCGTCTGCGCGCGCGCGAGCGCGGTTCCGCGCGTGT 5266

QY 121 TTCTGGGTGACCGGCTGATTTCTCAGCCCTTTCGCAATCCCTATATTCATCAACCAAC 180
DB 5267 TTCTGGGTGACCGGCTGATTTCTCAGCCCTTTCGCAATCCCTATATTCATCAACCAAC 5326

QY 181 CCCTTCGCGCCCGATGTCACCGCTGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 240
DB 5327 CCCTTCGCGCCCGATGTCACCGCTGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 5386

QY 241 CGACCACTCGGCTCGGCTGCGGTGACACGCGCCGCGCGCGCGCGCGTTCGCTCAGCGTGT 300
DB 5387 CGACCACTCGGCTCGGCTGCGGTGACACGCGCCGCGCGCGCGCGCGTTCGCTCAGCGTGT 5446

QY 301 AGACCTACACAGCTGGGCGCGCGGTAAACGCGGTGCTTCGCGCGCGATGACACCGCG 360
DB 5447 AGACCTACACAGCTGGGCGCGCGGTAAACGCGGTGCTTCGCGCGCGATGACACCGCG 5506

QY 361 CCAGTGCCTGATGTGACCTCCCGCGCGCATCTTTCGCGCGCGAGTATAACCTATCAACA 420
DB 5507 CCAGTGCCTGATGTGACCTCCCGCGCGCATCTTTCGCGCGCGAGTATAACCTATCAACA 5566

QY 421 TCTCCCTTACCTTTCGCTGGCCACCGGCAATAACCTGGTTCCTTATGCGCGCCCTCTT 480
DB 5567 TCTCCCTTACCTTTCGCTGGCCACCGGCAATAACCTGGTTCCTTATGCGCGCCCTCTT 5626

QY 481 AGTCCGCTTTTACCCCTTCAGACGCGCACCAATACCATATATATGCGCACCGAAGCTTCT 540
DB 5627 AGTCCGCTTTTACCCCTTCAGACGCGCACCAATACCATATATATGCGCACCGAAGCTTCT 5686
```

```
QY 1621 CTCGCGGTAGCTCTCTTTCTGGAGGAGCAGCAACTAAAGCGGTACCCCTTATAT 1680
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6767 CTCGCGGTAGCTCTCTTTCTGGAGGAGCAGCAACTAAAGCGGTACCCCTTATAT 6826
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1681 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATGCGCGGACCGGTCGCT 1740
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6827 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATGCGCGGACCGGTCGCT 6886
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1741 ATTTCACCTTACACCACTAGCTGGTGTCTGGTCCGCTCCATTTCTGCGGTTGCCGTT 1800
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6887 ATTTCACCTTACACCACTAGCTGGTGTCTGGTCCGCTCCATTTCTGCGGTTGCCGTT 6946
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCCCGGCC 1860
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6947 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCCCGGCC 7006
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTGCAGCCCCCTTGGCTTCAGGGCTGCGCTTTC 1920
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7007 CATACTTTTGATGATTTCTGCCAGAGTGCAGCCCCCTTGGCTTCAGGGCTGCGCTTTC 7066
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1921 CAGTCTACTGTGCTGAGCTTCAGGCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7067 CAGTCTACTGTGCTGAGCTTCAGGCTTAAGATGAAGTGGTAAACTCGGAGTTG 7126
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1981 TAG 1983
| | |
Db 7127 TAG 7129
| | |

RESULT 15
US-09-553-427-6
; Sequence 6, Application US/09553427
; Patent No. 6379891
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/553,427
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0183.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HEV - Burma strain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..5106
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5147..7126
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5106..5474
; US-09-553-427-6

Query Match 100.0%; Score 1983; DB 4; Length 7195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCCCTCGGCCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCGCGGCCA 60
Db 5147 ATGGCGCCCTCGGCCTATTTTGTGCTGCTCCTCATGTTTTCCTATGCTGCGCGGCCA 5206
QY 61 CGCGCGGTCAGCGCTCTGCGCGCGCGCTGCTGGGCGGCGCAGCGCGGTTCCGCGCGTGT 120
Db 5207 CGCGCGGTCAGCGCTCTGCGCGCGCGCTGCTGGGCGGCGCAGCGCGGTTCCGCGCGTGT 5266
QY 121 TTCTGGGCTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTTCATCCACCAAC 180
Db 5267 TTCTGGGCTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTTCATCCACCAAC 5326
QY 181 CCCTTCGCGCCCGATGTACACGCTGCGCGCGCGGCTGGACCTCGTGTTCGCAACCCGCC 240
Db 5327 CCCTTCGCGCCCGATGTACACGCTGCGCGCGCGGCTGGACCTCGTGTTCGCAACCCGCC 5386
QY 241 CGACCACTCGGCTCGCTGCGGTCGCTGACCGCGCCAGCGCGCGGTTGCTTCAGCTGCT 300
Db 5387 CGACCACTCGGCTCGCTGCGGTCGCTGACCGCGCCAGCGCGCGGTTGCTTCAGCTGCT 5446
QY 301 AGACTTACACAGCTGGGCGCGCGCTAACCGGCTGCTCCGCGCCATGACACCCCG 360
Db 5447 AGACTTACACAGCTGGGCGCGCGCTAACCGGCTGCTCCGCGCCATGACACCCCG 5506
QY 361 CCAGTGCCTGATGTGCACTCCCGCGCGCGCACTTTTGGCGCGGAGTATAAATCAACA 420
Db 5507 CCAGTGCCTGATGTGCACTCCCGCGCGCGCACTTTTGGCGCGGAGTATAAATCAACA 5566
QY 421 TCTCCCTTACCTCTTCGCTGGCCACCGGCACTAACCTGCTTTTATGCGCGCCCTCT 480
Db 5567 TCTCCCTTACCTCTTCGCTGGCCACCGGCACTAACCTGCTTTTATGCGCGCCCTCT 5626
```

QY 481 AGTCGCTTTTACCCCTTCAGACGGCACAATACCATATATATATGCGCACGGAAGCTTCT 540
Db
QY 5627 AGTCGCTTTTACCCCTTCAGACGGCACAATACCATATATATATGCGCACGGAAGCTTCT 5686
Db
QY 541 AATATATGCCAGTACCGGGTTGCCGTGCACAAATCCGTTACCGCCCGCTGTGTCCCAAT 600
Db
QY 5687 AATATATGCCAGTACCGGGTTGCCGTGCACAAATCCGTTACCGCCCGCTGTGTCCCAAT 5746
QY 601 GCTGTGCGGGTTACGCGCATCTCCATCTCATTTCTGCGCACAGACCAACCAACCCCGAGC 660
Db
QY 5747 GCTGTGCGGGTTACGCGCATCTCCATCTCATTTCTGCGCACAGACCAACCAACCCCGAGC 5806
QY 661 TCCGTTGATATGAATTAATTAACCTCGACGATGTTTCTGATTTTCTGATTCACAGCCCGGCATA 720
Db
QY 5807 TCCGTTGATATGAATTAATTAACCTCGACGATGTTTCTGATTTTCTGATTCACAGCCCGGCATA 5866
QY 721 GCTCTGAGCTTGATGCCAAGTGAAGCGCTACATATGTTAAACAAAGCTGGCGCTCC 780
Db
QY 5867 GCTCTGAGCTTGATGCCAAGTGAAGCGCTACATATGTTAAACAAAGCTGGCGCTCC 5926
QY 781 GTCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGCTCTGTTATGCTTTGCATA 840
Db
QY 5927 GTCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGCTCTGTTATGCTTTGCATA 5986
QY 841 CATGGCTCACTCGTAAATTCCTATATACTAATACACCTTATACCGGTGCCCTCGGGCTGTTG 900
Db
QY 5987 CATGGCTCACTCGTAAATTCCTATATACTAATACACCTTATACCGGTGCCCTCGGGCTGTTG 6046
QY 901 GACTTTGCGCTTGAGCTTGAGTTTGCACCTTACCCCGGTAAACCAATACGCGGTC 960
Db
QY 6047 GACTTTGCGCTTGAGCTTGAGTTTGCACCTTACCCCGGTAAACCAATACGCGGTC 6106
QY 961 TCCCGTTATTCAGCACTGCTCCACCGCTTCGTCGCGGTGCGACGCGGCTGCGGAG 1020
Db
QY 6107 TCCCGTTATTCAGCACTGCTCCACCGCTTCGTCGCGGTGCGACGCGGCTGCGGAG 6166
QY 1021 CTCACCAACACGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
Db
QY 6167 CTCACCAACACGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 6226
QY 1081 GTCGCTGAGATCGGCGCGGGATAGCCCTCACCTGTTCAACCTTCTGCTGACACTCTGCTT 1140
Db
QY 6227 GTCGCTGAGATCGGCGCGGGATAGCCCTCACCTGTTCAACCTTCTGCTGACACTCTGCTT 6286
QY 1141 GCGGGCTGCGGACAGAAATTTGATTTCTGCTGGCTGGTGGCCAGCTGTTCTACTCCGCTCC 1200
Db
QY 6287 GCGGGCTGCGGACAGAAATTTGATTTCTGCTGGCTGGTGGCCAGCTGTTCTACTCCGCTCC 6346
QY 1201 GTTGTCTCAGCCAATGGCGCGGACTGTTTAACTTGTATATACATCTGTAGAGAATGCTCAG 1260
Db
QY 6347 GTTGTCTCAGCCAATGGCGCGGACTGTTTAACTTGTATATACATCTGTAGAGAATGCTCAG 6406
QY 1261 CAGGATAAGGGTATTGGAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGGTTATT 1320
Db
QY 6407 CAGGATAAGGGTATTGGAATCCCGCATGACATTTGACCTCGGAGAAATCTCGTGTGGTTATT 6466
QY 1321 CAGGATTATGATAACCAACATGAACAGATTCGCGCGAGCGCTTCTCAGCCCATCGCGC 1380
Db
QY 6467 CAGGATTATGATAACCAACATGAACAGATTCGCGCGAGCGCTTCTCAGCCCATCGCGC 6526
QY 1381 CTTTCTCTGCTCTCGAGCTAATGATGTCTTTGGCTCTCTCTCAGCCGCTGCGGAGTAT 1440
Db
QY 6527 CTTTCTCTGCTCTCGAGCTAATGATGTCTTTGGCTCTCTCTCAGCCGCTGCGGAGTAT 6586
QY 1441 GACCACTCAGCTTATGGCTCTTGCAGTGGCGGAGTTTATGTTTCTGACTCTGAGCTTG 1500
Db
QY 6587 GACCACTCAGCTTATGGCTCTTGCAGTGGCGGAGTTTATGTTTCTGACTCTGAGCTTG 6646
QY 1501 GTTAATGTTGCGACCGCGCGGCGTTTGGCGGTTCGCTCGATTTGGACCAAGGTACCA 1560
Db
QY 6647 GTTAATGTTGCGACCGCGCGGCGGTTTGGCGGTTCGCTCGATTTGGACCAAGGTACCA 6706
QY 1561 CTTGACGGTGGCCCTCTCCACCAATCCAGCAGTACTCGAAGACCTTCTTTTGTCTGCGC 1620

Db 6707 CTTGACGGTGGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTGCGC 6766
QY 1621 CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGACAACTAAAGCCGGGTACCTTTATAAT 1680
Db 6767 CTCGCGGTAAGCTCTCTTTCTGGGAGGAGGACAACTAAAGCCGGGTACCTTTATAAT 6826
QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCGCGCGGACCGGGTCGCT 1740
Db 6827 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCGCGCGGACCGGGTCGCT 6886
QY 1741 ATTTCACCTTACACCACTAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 6887 ATTTCACCTTACACCACTAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6946
QY 1801 TTAGCCCCCCTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCGCGCGCC 1860
Db 6947 TTAGCCCCCCTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCGCGCGCC 7006
QY 1861 CATACTTTTGTGATTTCTGCGCAGAGTGGCGCGCGCTTGGCCTTCAGGGCTGCGCTTTC 1920
Db 7007 CATACTTTTGTGATTTCTGCGCAGAGTGGCGCGCGCTTGGCCTTCAGGGCTGCGCTTTC 7066
QY 1921 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAAGATGAAGGTGGGTAAACTCGGGAGTTG 1980
Db 7067 CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAAGATGAAGGTGGGTAAACTCGGGAGTTG 7126
QY 1981 TAG 1983
Db 7127 TAG 7129

Search completed: August 21, 2004, 18:10:41

Job time : 159.606 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 15:15:13 ; Search time 886.671 Seconds
(without alignments)
10985.769 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129

Perfect score: 1983
Sequence: 1 ATGCGCGCTCGGCTATT...GTAAACTCGGAGTGTAG 1983

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3228839 seqs, 245606551 residues

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubnpa/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/1/pubnpa/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/1/pubnpa/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/1/pubnpa/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/1/pubnpa/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/1/pubnpa/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/1/pubnpa/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/1/pubnpa/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/1/pubnpa/US09A_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/1/pubnpa/US09B_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/1/pubnpa/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/1/pubnpa/US09_NEW_PUB.seq:
- 13: /cgn2_6/ptodata/1/pubnpa/US09_NEW_PUB.seq2:
- 14: /cgn2_6/ptodata/1/pubnpa/US10A_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/1/pubnpa/US10B_PUBCOMB.seq:
- 16: /cgn2_6/ptodata/1/pubnpa/US10C_PUBCOMB.seq:
- 17: /cgn2_6/ptodata/1/pubnpa/US10_NEW_PUB.seq:
- 18: /cgn2_6/ptodata/1/pubnpa/US60_NEW_PUB.seq:
- 19: /cgn2_6/ptodata/1/pubnpa/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1983	100.0	2049	9	US-09-769-066-1
2	1983	100.0	2049	15	US-10-165-868-3
3	1983	100.0	2094	15	US-10-165-868-1
4	1983	100.0	7195	10	US-09-851-410-6
5	1805.4	91.0	1990	13	US-10-381-770-5
6	1793.4	90.4	1989	13	US-10-381-770-6
7	1647	83.1	1647	9	US-09-769-066-3
8	1384.6	69.8	2058	9	US-09-769-066-2
9	1368.6	69.0	2055	15	US-10-165-868-4
10	1368.6	69.0	2100	15	US-10-165-868-2
11	1368.6	69.0	7171	10	US-09-851-410-10
12	1368.6	69.0	7180	17	US-10-239-090A-49
13	1367.6	69.0	7230	17	US-10-239-090A-48
14	1347.6	68.0	7228	17	US-10-239-090A-47

15	1346.8	67.9	7233	17	US-10-239-090A-45
16	1346	67.9	7234	17	US-10-239-090A-11
17	1344.4	67.8	7239	17	US-10-239-090A-46
18	1343.6	67.8	7256	17	US-10-239-090A-44
19	1339.6	67.6	7277	10	US-09-468-147-164
20	1339.6	67.6	7277	10	US-09-468-147-165
21	1339.6	67.6	7277	13	US-10-319-745-164
22	1339.6	67.6	7277	13	US-10-319-745-165
23	1323.2	66.7	7202	10	US-09-468-147-89
24	1323.2	66.7	7202	10	US-09-468-147-90
25	1323.2	66.7	7202	13	US-10-319-745-89
26	1323.2	66.7	7202	13	US-10-319-745-90
27	1311	66.1	1311	15	US-10-165-868-9
28	1106.2	55.8	1647	9	US-09-769-066-4
29	984	49.6	984	9	US-09-769-066-5
30	984	49.6	984	15	US-10-165-868-7
31	880.6	44.4	1308	15	US-10-165-868-10
32	675.6	34.1	1020	10	US-09-468-147-129
33	675.6	34.1	1020	13	US-10-319-745-129
34	673.6	34.0	984	9	US-09-769-066-6
35	672.2	33.9	981	15	US-10-165-868-8
36	658.8	33.2	1389	10	US-09-468-147-193
37	658.8	33.2	1389	10	US-09-468-147-196
38	658.8	33.2	1389	13	US-10-319-745-193
39	658.8	33.2	1389	13	US-10-319-745-196
40	652.6	32.9	1026	10	US-09-468-147-192
41	652.6	32.9	1026	10	US-09-468-147-195
42	652.6	32.9	1026	13	US-10-319-745-192
43	652.6	32.9	1026	13	US-10-319-745-195
44	583.8	29.4	899	10	US-09-468-147-24
45	583.8	29.4	899	13	US-10-319-745-24

ALIGNMENTS

RESULT 1

US-09-769-066-1
; Sequence 1, Application US/09769066
; Patent No. US20020107360A1

GENERAL INFORMATION:

APPLICANT: Fuerst, Thomas R.
McAttee, C. Patrick
Yarborough, Patrice O.
Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Ave., Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/769,066
FILING DATE: 24-Jan-2001
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/542,634
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960

Sequence 45, Appl
Sequence 11, Appl
Sequence 46, Appl
Sequence 44, Appl
Sequence 164, Appl
Sequence 165, Appl
Sequence 164, Appl
Sequence 89, Appl
Sequence 90, Appl
Sequence 89, Appl
Sequence 90, Appl
Sequence 9, Appl
Sequence 4, Appl
Sequence 5, Appl
Sequence 7, Appl
Sequence 10, Appl
Sequence 129, Appl
Sequence 129, Appl
Sequence 6, Appl
Sequence 8, Appl
Sequence 193, Appl
Sequence 196, Appl
Sequence 193, Appl
Sequence 196, Appl
Sequence 192, Appl
Sequence 195, Appl
Sequence 192, Appl
Sequence 195, Appl
Sequence 24, Appl
Sequence 24, Appl

```
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2049 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: Hepatitis E Virus (Burma strain)
;     TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     ORIGINAL SOURCE:
;     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-769-066-1

Query Match      100.0%; Score 1983; DB 9; Length 2049;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCTCGGCTATTTGTTGCTCTCATGTTTGTGCTATGCTGCGCGCCA 60
DB 1 ATGCGCCTCGGCTATTTGTTGCTCTCATGTTTGTGCTATGCTGCGCGCCA 60
QY 61 CCGCCCGGTCAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGT 120
DB 61 CCGCCCGGTCAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGT 120
QY 121 TTCTGGGGTACCGGCTGATTCTCAGCCCTTCGCAATCCCTCATATTATTCATCCAAAC 180
DB 121 TTCTGGGGTACCGGCTGATTCTCAGCCCTTCGCAATCCCTCATATTATTCATCCAAAC 180
QY 181 CCTTCGCCCCGATGTCACCGCTGCGCGCGCGCGCGCGCGCGCGT 240
DB 181 CCTTCGCCCCGATGTCACCGCTGCGCGCGCGCGCGCGCGCGCGT 240
QY 241 CGACCACTCGGCTCCGCTTGGCGTGACAGGCCAGCGCGCGCGCGCGT 300
DB 241 CGACCACTCGGCTCCGCTTGGCGTGACAGGCCAGCGCGCGCGCGCGT 300
QY 301 AGACTTACACAGCTGGGCGCGCGCGCTAACCGCGGTGCGTCCGCGCCATGACACCGCG 360
DB 301 AGACTTACACAGCTGGGCGCGCGCGCTAACCGCGGTGCGTCCGCGCCATGACACCGCG 360
QY 361 CCAGTGCCTGATGTCAGCTCCCGCGCGCGCGCGCGCGCGCGT 420
DB 361 CCAGTGCCTGATGTCAGCTCCCGCGCGCGCGCGCGCGCGCGT 420
QY 421 TCTCCCTTACCTCTTCCGTGCGCACCGCGCTAACCTGCTGTTCTTTATGCGCGCGCTCTT 480
DB 421 TCTCCCTTACCTCTTCCGTGCGCACCGCGCTAACCTGCTGTTCTTTATGCGCGCGCTCTT 480
QY 481 AGTCCGCTTTTACCGCTTCAGGACGGCACCAATACCCATATATATGCGCGCGAGCTTCT 540
DB 481 AGTCCGCTTTTACCGCTTCAGGACGGCACCAATACCCATATATATGCGCGCGAGCTTCT 540
QY 541 AATATGCGCGTACCGGTTGCGCGTGCACCAATCCGTTACCGCGCGTTCGCCAAT 600
DB 541 AATATGCGCGTACCGGTTGCGCGTGCACCAATCCGTTACCGCGCGTTCGCCAAT 600
QY 601 GCTGTCGCGGTACCGCATCTCCATCTCACTTCTGCGCACAGCACCAACCGCGAG 660
DB 601 GCTGTCGCGGTACCGCATCTCCATCTCACTTCTGCGCACAGCACCAACCGCGAG 660
QY 661 TCCGTTGATATGAATTCATACCTCGAGCGGATGTTGTTATGTCAGCGCGGATA 720
DB 661 TCCGTTGATATGAATTCATACCTCGAGCGGATGTTGTTATGTCAGCGCGGATA 720
QY 721 GCCTCTGAGCTTGATCCCAAGTACGCGCTTACATATGCTTACCAAGGCTGCGCTCC 780
DB 721 GCCTCTGAGCTTGATCCCAAGTACGCGCTTACATATGCTTACCAAGGCTGCGCTCC 780
QY 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGCTGTTGTTATGCTTGCATA 840
DB 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGCTACCTCTGCTGTTGTTATGCTTGCATA 840
```

```
QY 841 CATGGCTCACTCGTAATTCCTATCTAATAACACCTATACCGGTGCGCTCGGCTGTTG 900
DB 841 CATGGCTCACTCGTAATTCCTATCTAATAACACCTATACCGGTGCGCTCGGCTGTTG 900
QY 901 GACTTTGCCCTTCAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATACGCGGTC 960
DB 901 GACTTTGCCCTTCAGCTTGAGTTTCGCAACCTTTACCCCGGTAAACCAATACGCGGTC 960
QY 961 TCCCGTTATTCAGCACTGCTGCCACCGCTTTGTCGCGGTGCGGACGCGGACTGCCGAG 1020
DB 961 TCCCGTTATTCAGCACTGCTGCCACCGCTTTGTCGCGGTGCGGACGCGGACTGCCGAG 1020
QY 1021 CTCACCAACGAGCTGCTACCGCTTTATGAAGACCTCTATTCTACTAGTACTAATGGT 1080
DB 1021 CTCACCAACGAGCTGCTACCGCTTTATGAAGACCTCTATTCTACTAGTACTAATGGT 1080
QY 1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
DB 1081 GTCGGTGAGATCGGCGCGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
QY 1141 GCGCGCTCGCGACAGATTGATTTCGTCGGCTGGTGGCGAGCTGTTACTCCCGTCCC 1200
DB 1141 GCGCGCTCGCGACAGATTGATTTCGTCGGCTGGTGGCGAGCTGTTACTCCCGTCCC 1200
QY 1201 GTTGTCTCAGCCAAATGGCGAGCGGCTGTTAAAGTTGTATACATCTCTAGAGATGCTCAG 1260
DB 1201 GTTGTCTCAGCCAAATGGCGAGCGGCTGTTAAAGTTGTATACATCTCTAGAGATGCTCAG 1260
QY 1261 CAGGATAAGGGTATTGCAATCCGATGACATTGACCTCGGAGAACTCTGCTGGTATT 1320
DB 1261 CAGGATAAGGGTATTGCAATCCGATGACATTGACCTCGGAGAACTCTGCTGGTATT 1320
QY 1321 CAGGATTATGATAACCAATGAAACAGATCGGCGAGCGCTTCTCAGCCCATCGCG 1380
DB 1321 CAGGATTATGATAACCAATGAAACAGATCGGCGAGCGCTTCTCAGCCCATCGCG 1380
QY 1381 CTTTCTCTGCTTTCGAGCTAAATGATGCTGCTTGGCTCTCTCTCAGCGCTCGGAGTAT 1440
DB 1381 CTTTCTCTGCTTTCGAGCTAAATGATGCTGCTTGGCTCTCTCTCAGCGCTCGGAGTAT 1440
QY 1441 GACCACTCACTTATGGCTCTTTCGACTGCGCCAGTTTATGTTCTGACTCTGCTGCTTG 1500
DB 1441 GACCACTCACTTATGGCTCTTTCGACTGCGCCAGTTTATGTTCTGACTCTGCTGCTTG 1500
QY 1501 GTTATGTTGCGACCGCGCGCGCGCTGTCGCGGTGCGTCTGATGGAACCAAGTCA 1560
DB 1501 GTTATGTTGCGACCGCGCGCGCGCTGTCGCGGTGCGTCTGATGGAACCAAGTCA 1560
QY 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAGACCTTCTTGTCTGCGG 1620
DB 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAGACCTTCTTGTCTGCGG 1620
QY 1621 CTCGCGGTAAAGCTCTTCTTTCGAGGAGCGAGCACAACTAAAGCGGGTACCCCTTATAAT 1680
DB 1621 CTCGCGGTAAAGCTCTTCTTTCGAGGAGCGAGCACAACTAAAGCGGGTACCCCTTATAAT 1680
QY 1681 TATAACCACTGCTAGCGCACTGTTGTCGAGATGCGCGCGGAGACCGGGTCTGCT 1740
DB 1681 TATAACCACTGCTAGCGCACTGTTGTCGAGATGCGCGCGGAGACCGGGTCTGCT 1740
QY 1741 ATTTCCACTTACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB 1741 ATTTCCACTTACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
QY 1801 TTAGCCCGCACTCTGGCTAGCACTGCTGAGGATACCTTGGACTACCTTGCCTGCGCGCC 1860
DB 1801 TTAGCCCGCACTCTGGCTAGCACTGCTGAGGATACCTTGGACTACCTTGCCTGCGCGCC 1860
QY 1861 CATACTTTTATGATGTTCTGCGCGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
DB 1861 CATACTTTTATGATGTTCTGCGCGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
QY 1921 CAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
```


Db 1921 CAGTCTACTGCTGCTGAGCTTACGCGCTTAAAGATGAGGTGGTAAACTCGGGAGTTG 1980
 Qy 1981 TAG 1983
 Db 1981 TAG 1983

RESULT 2
 US-10-165-868-3
 ; Sequence 3, Application US/10165868
 ; Publication No. US20030143241A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reyes, Gregory R.
 ; Bradley, Daniel W.
 ; Twu, Jr-Shin
 ; Purdy, Michael A.
 ; Tam, Albert W.
 ; Krawczynski, Krzysztof Z.
 ; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: 350 Cambridge Avenue, Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/165,868
 ; FILING DATE: 06-Jun-2002
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/870,985A
 ; FILING DATE: 20-APRIL-1992
 ; APPLICATION NUMBER: US 822,335
 ; FILING DATE: 17-JAN-1992
 ; APPLICATION NUMBER: US 505,888
 ; FILING DATE: 05-APRIL-1990
 ; APPLICATION NUMBER: US 420,921
 ; FILING DATE: 13-OCTOBER-1989
 ; APPLICATION NUMBER: US 367,486
 ; FILING DATE: 16-JUNE-1989
 ; APPLICATION NUMBER: US 336,672
 ; FILING DATE: 11-APRIL-1989
 ; APPLICATION NUMBER: US 208,997
 ; FILING DATE: 17-JUNE-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fabian, Gary R.
 ; REGISTRATION NUMBER: 33,875
 ; REFERENCE/DOCKET NUMBER: 4600-0093.30
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 324-0880
 ; TELEFAX: (415) 324-0960
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2049 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cdna to mRNA
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 7
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 ; US-10-165-868-3

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGCCCTCGGCTATTTTGTGCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
 Db 1 ATGCGCCCTCGGCTATTTTGTGCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60

Qy 61 CGCCCGGTCAGCGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 Db 61 CGCCCGGTCAGCGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120

Qy 121 TTCTGGGGTGACCGGGTTGATTTCTCAGCCCTTTCGCAATCCCTATATTTATCAACCAAC 180
 Db 121 TTCTGGGGTGACCGGGTTGATTTCTCAGCCCTTTCGCAATCCCTATATTTATCAACCAAC 180

Qy 181 CCCTTCGCCCCGATGTCAACCGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGG 240
 Db 181 CCCTTCGCCCCGATGTCAACCGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGG 240

Qy 241 CGACCACTCGGCTCGGCTTGGGCGTGACAGGCGCGGCGGCGGCGGCGGCGGCGG 300
 Db 241 CGACCACTCGGCTCGGCTTGGGCGTGACAGGCGCGGCGGCGGCGGCGGCGGCGG 300

Qy 301 AGACCTTACACAGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 Db 301 AGACCTTACACAGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360

Qy 361 CAGTGCCTGATGCTGACTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 Db 361 CAGTGCCTGATGCTGACTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420

Qy 421 TCTCCCTTACCTTTCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 480
 Db 421 TCTCCCTTACCTTTCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 480

Qy 481 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATATATGCGGCGGCGGCT 540
 Db 481 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATATATGCGGCGGCGGCT 540

Qy 541 AATTATGCCAGTACCGGGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
 Db 541 AATTATGCCAGTACCGGGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600

Qy 601 GCTGTGCGGGTTACCGCATCTCCATCTCATTTGCGGCGGCGGCGGCGGCGGCGG 660
 Db 601 GCTGTGCGGGTTACCGCATCTCCATCTCATTTGCGGCGGCGGCGGCGGCGGCGG 660

Qy 661 TCCGTTGATATGAATTAATTAACCTCGACGGATGTTTCTGTTTATGTCAGCGCGG 720
 Db 661 TCCGTTGATATGAATTAATTAACCTCGACGGATGTTTCTGTTTATGTCAGCGCGG 720

Qy 721 GCTCTGAGCTTGTGATTCCTAATGAGCGGCTACATCTGTAACCAAGCGTGGCGTCC 780
 Db 721 GCTCTGAGCTTGTGATTCCTAATGAGCGGCTACATCTGTAACCAAGCGTGGCGTCC 780

Qy 781 GTCGAGACCTTCGGGTTGCTGAGGAGGAGGCTACCTCTGTTTATGTTTATGTTG 840
 Db 781 GTCGAGACCTTCGGGTTGCTGAGGAGGAGGCTACCTCTGTTTATGTTTATGTTG 840

Qy 841 CATGGCTCACTCGTAAATTCCTATATACACCTTATACCGTGGCTCGGCGGTTG 900
 Db 841 CATGGCTCACTCGTAAATTCCTATATACACCTTATACCGTGGCTCGGCGGTTG 900

Qy 901 GACTTTGCCCTTGAAGTTTTCGCACTTACCGCGGCGGCGGCGGCGGCGGCGGCGG 960
 Db 901 GACTTTGCCCTTGAAGTTTTCGCACTTACCGCGGCGGCGGCGGCGGCGGCGGCGG 960

Qy 961 TCCGTTATTCAGAGCTGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 Db 961 TCCGTTATTCAGAGCTGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020

Qy 1021 CTCAACACCGGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTATGTT 1080
 Db 1021 CTCAACACCGGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTATGTT 1080

181 CCCTTCGCCCCCGATGTCACCGCTGCGGCGCGGCTGGACCTCGTGTTCGCCAACCCGCG 240
Db |||||
226 CCCTTCGCCCCCGATGTCACCGCTGCGGCGCGGCTGGACCTCGTGTTCGCCAACCCGCG 285
Qy |||||
241 CGACCACTCGGCTCGCTTGGGCTGACCGAGCCAGCGCCGCGCTTGCCTCAGCTCGT 300
Db |||||
286 CGACCACTCGGCTCGCTTGGGCTGACCGAGCCAGCGCCGCGCTTGCCTCAGCTCGT 345
Qy |||||
301 AGACCTACACAGCTGGGCGCGGCGCTTAACCGGCTGCTCCGGCCATGACACCCG 360
Db |||||
346 AGACCTACACAGCTGGGCGCGGCGCTTAACCGGCTGCTCCGGCCATGACACCCG 405
Qy |||||
361 CCAGTGCCTGATGTCGACTCCGCGCGGCATCTTCGCGCGGCGAGTAACTATCAACA 420
Db |||||
406 CCAGTGCCTGATGTCGACTCCGCGCGGCGATCTTCGCGCGGCGAGTAACTATCAACA 465
Qy |||||
421 TCTCCCTTACCTTTCGCTGGCCACCGGCCTAACCTGCTTCTTTATGCCGCCCTCTT 480
Db |||||
466 TCTCCCTTACCTTTCGCTGGCCACCGGCCTAACCTGCTTCTTTATGCCGCCCTCTT 525
Qy |||||
481 AGTCCGCTTTTACCCCTTCAGACCGCACCAATACCCATATATGCGCCAGGAAGTTCT 540
Db |||||
526 AGTCCGCTTTTACCCCTTCAGACCGCACCAATACCCATATATGCGCCAGGAAGTTCT 585
Qy |||||
541 AATTATGCCAGTACCGGCTTGGCGGTGCGCAACAATCCGTTACCGCGCGCTGCTCCCAAT 600
Db |||||
586 AATTATGCCAGTACCGGCTTGGCGGTGCGCAACAATCCGTTACCGCGCGCTGCTCCCAAT 645
Qy |||||
601 GCTGTCGGGGTTACGCCATCTCCATCTCATCTTGGCCACAGACCAACCAACCCGAGC 660
Db |||||
646 GCTGTCGGGGTTACGCCATCTCCATCTCATCTTGGCCACAGACCAACCAACCCGAGC 705
Qy |||||
661 TCCGTTGATCAATCAATACCTCGAGGATGTCGATTTTATGTCAGCCCGGCATA 720
Db |||||
706 TCCGTTGATCAATCAATACCTCGAGGATGTCGATTTTATGTCAGCCCGGCATA 765
Qy |||||
721 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
Db |||||
766 GCCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 825
Qy |||||
826 GTCGAGACCTCTGGGCTGGCTGAGGAGGAGCTACCTCTGGTCTTGTATGCTTTGCATA 885
Db |||||
841 CATGCTCACTCGTAAATTCCTATCTACTAATACCCCTATACCGTGCCCTCGGCTGTG 900
Qy |||||
886 CATGCTCACTCGTAAATTCCTATCTACTAATACCCCTATACCGTGCCCTCGGCTGTG 945
Db |||||
901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATAACGCGGTC 960
Db |||||
946 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATAACGCGGTC 1005
Qy |||||
961 TCCCGTTATTCAGACCTGCTGCGCACCGCTTCGTCGCGTGGGAGCGGACTCGCGAG 1020
Db |||||
1006 TCCCGTTATTCAGACCTGCTGCGCACCGCTTCGTCGCGTGGGAGCGGACTCGCGAG 1065
Qy |||||
1021 CTCACCAACCGCTGCTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGTT 1080
Db |||||
1066 CTCACCAACCGCTGCTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGTT 1125
Qy |||||
1081 GTCGCTGAGATCGGCGCGGATAGCCCTCACCTCTGTTCAACCTTGTGACACTCTGCTT 1140
Db |||||
1126 GTCGCTGAGATCGGCGCGGATAGCCCTCACCTCTGTTCAACCTTGTGACACTCTGCTT 1185
Qy |||||
1141 GCGCGCTTCGCGACAGAAATGATTTGTCGGCTGTCGCGGCTTCTACTCCGCTCC 1200
Db |||||
1186 GCGCGCTTCGCGACAGAAATGATTTGTCGGCTGTCGCGGCTTCTACTCCGCTCC 1245
Qy |||||
1201 GTTGTCTAGCAATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAACTCTCAG 1260
Db |||||
1246 GTTGTCTAGCAATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAACTCTCAG 1305

1261 CAGGATAAGGGTATTGCAATCCGCGATGACATTGACCTCGGAGAACTCGTGTGGTTATT 1320
Db |||||
1306 CAGGATAAGGGTATTGCAATCCGCGATGACATTGACCTCGGAGAACTCGTGTGGTTATT 1365
Qy |||||
1321 CAGGATTATGATAACCAACATGAACAAGATCGGCGGAGCCCTTCTCCAGCCCCATCGCGC 1380
Db |||||
1366 CAGGATTATGATAACCAACATGAACAAGATCGGCGGAGCCCTTCTCCAGCCCCATCGCGC 1425
Qy |||||
1381 CTTTCTCTGTCTCTGAGCTAATGATGCTTTGGCTCTCTCTCACCGTGCAGATAT 1440
Db |||||
1426 CTTTCTCTGTCTCTGAGCTAATGATGCTTTGGCTCTCTCTCACCGTGCAGATAT 1485
Qy |||||
1441 GACCACTCACATTAAGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db |||||
1486 GACCACTCACATTAAGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1545
Qy |||||
1501 GTTAATGTTGCGACCGGCGCGAGGCGCTTGCCCGTCTCGATTGGACCAAGGTCACA 1560
Db |||||
1546 GTTAATGTTGCGACCGGCGCGAGGCGCTTGCCCGTCTCGATTGGACCAAGGTCACA 1605
Qy |||||
1561 CTTGAGCGTTCGCCCTCTCTCCATTCAGCAGTACTCGAAGACCTTTCTTCTGCTGCGC 1620
Db |||||
1606 CTTGAGCGTTCGCCCTCTCTCCATTCAGCAGTACTCGAAGACCTTTCTTCTGCTGCGC 1665
Qy |||||
1621 CTCGCGGTAACTCTCTTTCTGGGAGGAGGAGCAACAATAAGCGGGTACCTTATAAT 1680
Db |||||
1666 CTCGCGGTAACTCTCTTTCTGGGAGGAGGAGCAACAATAAGCGGGTACCTTATAAT 1725
Qy |||||
1681 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATGCCCGCGGCGACCGGGTCGCT 1740
Db |||||
1726 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATGCCCGCGGCGACCGGGTCGCT 1785
Qy |||||
1741 ATTTCACCTTACACCACTGTAGCGTGGTGTGCTGCTTCCATTTCTGCGGTTCGCGTT 1800
Db |||||
1786 ATTTCACCTTACACCACTGTAGCGTGGTGTGCTGCTTCCATTTCTGCGGTTCGCGTT 1845
Qy |||||
1801 TTAGCCCCCACTCTGCGTAGCATTTGCTGAGGATACCTTGGACTACCTTCCCGCGCGC 1860
Db |||||
1846 TTAGCCCCCACTCTGCGTAGCATTTGCTGAGGATACCTTGGACTACCTTCCCGCGCGC 1905
Qy |||||
1861 CATACTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCTTTCAGGGCTGCGCTTTC 1920
Db |||||
1906 CATACTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCTTTCAGGGCTGCGCTTTC 1965
Qy |||||
1921 CAGTCTACTGCTGCTGAGCTTACGCGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
Db |||||
1966 CAGTCTACTGCTGCTGAGCTTACGCGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 2025
Qy |||||
1981 TAG 1983
Db |||||
2026 TAG 2028

RESULT 4

US-09-851-410-6

; Sequence 6, Application US/09851410

; Publication No. US20030124510A1

; GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory R

; Yarbrough, Patrice O

; Bradley, Daniel W

; Krawczynski, Krzysztof Z

; Tam, Albert

; Fry, Kirk E

; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted

; No. US20030124510A1-A/No. US20030124510A1-B Hepatitis Viral

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates

; STREET: 350 Cambridge Avenue, Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

1141 GCGCGCTCCGACAGAAATGATTTCTGCGCTGGTGGCCAGCTGTTCTACTCCGTCCTC 1200
1141 GCGCGCTCCGACAGAAATGATTTCTGCGCTGGTGGCCAGCTGTTCTACTCCGTCCTC 1200
1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGATATACATCTGTAGAAATGCTCAG 1260
1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGATATACATCTGTAGAAATGCTCAG 1260
1261 CAGGATAGGATGATGCAATCCCGATGACATCTCGGAGAAATCTGTTGTTAT 1320
1261 CAGGATAGGATGATGCAATCCCGATGACATCTCGGAGAAATCTGTTGTTAT 1320
1321 CAGGATATGATAAACAATGAAACAAGATCGGCGGAGCTTCTCAGCCCATCGGC 1380
1321 CAGGATATGATAAACAATGAAACAAGATCGGCGGAGCTTCTCAGCCCATCGGC 1380
1381 CCTTTCTCTGCTTCCGAGCTAATGATGCTTTCGCTCTCTCAGCGCTCCGAGTAT 1440
1381 CCTTTCTCTGCTTCCGAGCTAATGATGCTTTCGCTCTCTCAGCGCTCCGAGTAT 1440
1441 GACCACTCCACTTATGCTTCTGACTGGCCCAAGTTTATGTTTCTGACTCTGAGACTTGG 1500
1441 GACCACTCCACTTATGCTTCTGACTGGCCCAAGTTTATGTTTCTGACTCTGAGACTTGG 1500
1501 GTTAATGTTGCGACCGCGCGCAGGCGGTTGCGCGCTGCTGATGGAAGTCA 1560
1501 GTTAATGTTGCGACCGCGCGCAGGCGGTTGCGCGCTGCTGATGGAAGTCA 1560
1561 CTTGAGGCTGCGCCCTCTCCACCATCCAGCTGCTCGAAGACTTCTTGTCTGCGG 1620
1561 CTTGAGGCTGCGCCCTCTCCACCATCCAGCTGCTCGAAGACTTCTTGTCTGCGG 1620
1621 CTCGCGGTAAAGCTCTCTTTCTGGGAGGCGACAACTAAAGCCGGTACCTTATAAT 1680
1621 CTCGCGGTAAAGCTCTCTTTCTGGGAGGCGAGGTACTTAAAGCCGGTACCTTATAAT 1680
1681 TATAACCACTCTAGTACCAACTGCTGTTGAGAAATCCGCTGGGCAATCGGTTGCT 1740
1681 TATAACCACTCTAGTACCAACTGCTGTTGAGAAATCCGCTGGGCAATCGGTTGCT 1740
1741 ATTTCCACTTACCACTAGCTGGGTTGCTGCTCCGCTCCCAATTTCTGCGGTTGCTGTT 1800
1741 ATTTCCACTTACCACTAGCTGGGTTGCTGCTCCGCTCCCAATTTCTGCGGTTGCTGTT 1800
1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTTGAGGATACCTTGAAGTACCTGCGCGCC 1860
1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTTGAGGATACCTTGAAGTACCTGCGCGCC 1860
1861 CATACTTTGATGATTTCTGCGAGAGTGGCGCCCTTGGCTTCAAGGCTGCGCTTTC 1920
1861 CATACTTTGATGATTTCTGCGAGAGTGGCGCCCTTGGCTTCAAGGCTGCGCTTTC 1920
1921 CAGTCTACTGCTGAGCTTCAAGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
1921 CAGTCTACTGCTGAGCTTCAAGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
1981 TAG 1983
1981 TAG 1983

Query Match 90.4%; Score 1793.4; DB 13; Length 1989;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 1871; Conservative 0; Mismatches 111; Indels 1; Gaps 1;

1 ATGCGCCCTCGGCTTATTTTGTGCTGCTCTCTCATGTTTTCCTATCTGTCGCCGCGCA 60
1 ATGCGCCCTCGGCTTATTTTGTGCTGCTCTCTCATGTTTTCCTATCTGTCGCCGCGCA 60
61 CCGCGCGGTGACCGCTCTGCGCGCGCTGTCGCGCGCGCGCGCGCTGTCGCCGCGGTGT 120
61 CCGCGCGGTGACCGCTCTGCGCGCGCTGTCGCGCGCGCGCGCGCTGTCGCCGCGGTGT 120
121 TTCTCGGGTGACCGGCTTGAATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
121 TTCTCGGGTGACCGGCTTGAATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
181 CCTTTCGCGCGGTGACCGCTCTGCGCGCGCTGTCGCGCGCGCGCGCTGTCGCCGCGGTGT 240
181 CCTTTCGCGCGGTGACCGCTCTGCGCGCGCTGTCGCGCGCGCGCGCTGTCGCCGCGGTGT 240
241 CGACCACTCGGCTTGGGCTGACAGGCGCGCGCGCTTTCGCTGCTGCTGCTGCTGCTGCT 300
241 CGACCACTCGGCTTGGGCTGACAGGCGCGCGCGCTTTCGCTGCTGCTGCTGCTGCTGCT 300
301 AGACCTTACCACTGCGGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCGCTGTCGCCGCGGTGT 360
301 AGACCTTACCACTGCGGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCGCTGTCGCCGCGGTGT 360
361 CGAGTCTGATGTCGCTGCGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCGCTTAAACCA 420
361 CGAGTCTGATGTCGCTGCGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCGCTTAAACCA 420
421 TCTCCCTTACCTTCCGCTGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCTTAAACCA 480
421 TCTCCCTTACCTTCCGCTGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCTTAAACCA 480
481 AGTCCGCTTTCACCTTTCAGGACCGCAATCAATCAATCAATCAATCAATCAATCAATCAAT 540
481 AGTCCGCTTTCACCTTTCAGGACCGCAATCAATCAATCAATCAATCAATCAATCAATCAAT 540
541 AATATGCCAGTACCGGCTTCCGCTGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCTTAAAC 600
541 AATATGCCAGTACCGGCTTCCGCTGCGCGCTTAAACCGCGCTGCTGCGCGCGCGCTTAAAC 600
601 GCTGTCGCGGTTACGCTTCCATCTCTTTCGCGCGCTTAAACCGCGCTGCTGCGCGCGCTTAAAC 660
601 GCTGTCGCGGTTACGCTTCCATCTCTTTCGCGCGCTTAAACCGCGCTGCTGCGCGCGCTTAAAC 660
661 TCCGTTGATGATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 720
661 TCCGTTGATGATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 720
721 GCCTCTGAGCTTGTATCCCAAGTGGCTTACACTATCGTAACCAAGGCTGGGCTTCC 780
721 GCCTCTGAGCTTGTATCCCAAGTGGCTTACACTATCGTAACCAAGGCTGGGCTTCC 780
781 GTCGAGACTCTCGGGTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
781 GTCGAGACTCTCGGGTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
841 CATGCTCACTCGTAAATTCCTATCACTAATCACTTACCGGCTGCTGCTGCTGCTGCTGCTGCTG 900
841 CATGCTCACTCGTAAATTCCTATCACTAATCACTTACCGGCTGCTGCTGCTGCTGCTGCTGCTG 900
901 GACTTTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 960

1141 GCGCGCTCCGACAGAAATGATTTCTGCGCTGGTGGCCAGCTGTTCTACTCCGTCCTC 1200
1141 GCGCGCTCCGACAGAAATGATTTCTGCGCTGGTGGCCAGCTGTTCTACTCCGTCCTC 1200
1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGATATACATCTGTAGAAATGCTCAG 1260
1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGATATACATCTGTAGAAATGCTCAG 1260
1261 CAGGATAGGATGATGCAATCCCGATGACATCTCGGAGAAATCTGTTGTTAT 1320
1261 CAGGATAGGATGATGCAATCCCGATGACATCTCGGAGAAATCTGTTGTTAT 1320
1321 CAGGATATGATAAACAATGAAACAAGATCGGCGGAGCTTCTCAGCCCATCGGC 1380
1321 CAGGATATGATAAACAATGAAACAAGATCGGCGGAGCTTCTCAGCCCATCGGC 1380
1381 CCTTTCTCTGCTTCCGAGCTAATGATGCTTTCGCTCTCTCAGCGCTCCGAGTAT 1440
1381 CCTTTCTCTGCTTCCGAGCTAATGATGCTTTCGCTCTCTCAGCGCTCCGAGTAT 1440
1441 GACCACTCCACTTATGCTTCTGACTGGCCCAAGTTTATGTTTCTGACTCTGAGACTTGG 1500
1441 GACCACTCCACTTATGCTTCTGACTGGCCCAAGTTTATGTTTCTGACTCTGAGACTTGG 1500
1501 GTTAATGTTGCGACCGCGCGCAGGCGGTTGCGCGCTGCTGATGGAAGTCA 1560
1501 GTTAATGTTGCGACCGCGCGCAGGCGGTTGCGCGCTGCTGATGGAAGTCA 1560
1561 CTTGAGGCTGCGCCCTCTCCACCATCCAGCTGCTCGAAGACTTCTTGTCTGCGG 1620
1561 CTTGAGGCTGCGCCCTCTCCACCATCCAGCTGCTCGAAGACTTCTTGTCTGCGG 1620
1621 CTCGCGGTAAAGCTCTCTTTCTGGGAGGCGACAACTAAAGCCGGTACCTTATAAT 1680
1621 CTCGCGGTAAAGCTCTCTTTCTGGGAGGCGAGGTACTTAAAGCCGGTACCTTATAAT 1680
1681 TATAACCACTCTAGTACCAACTGCTGTTGAGAAATCCGCTGGGCAATCGGTTGCT 1740
1681 TATAACCACTCTAGTACCAACTGCTGTTGAGAAATCCGCTGGGCAATCGGTTGCT 1740
1741 ATTTCCACTTACCACTAGCTGGGTTGCTGCTCCGCTCCCAATTTCTGCGGTTGCTGTT 1800
1741 ATTTCCACTTACCACTAGCTGGGTTGCTGCTCCGCTCCCAATTTCTGCGGTTGCTGTT 1800
1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTTGAGGATACCTTGAAGTACCTGCGCGCC 1860
1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTTGAGGATACCTTGAAGTACCTGCGCGCC 1860
1861 CATACTTTGATGATTTCTGCGAGAGTGGCGCCCTTGGCTTCAAGGCTGCGCTTTC 1920
1861 CATACTTTGATGATTTCTGCGAGAGTGGCGCCCTTGGCTTCAAGGCTGCGCTTTC 1920
1921 CAGTCTACTGCTGAGCTTCAAGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
1921 CAGTCTACTGCTGAGCTTCAAGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
1981 TAG 1983
1981 TAG 1983

US-10-381-770-6
Sequence 6, Application US/10381770
Publication No. US20040052813A1
GENERAL INFORMATION:
APPLICANT: YANG SHENG TANG company, Ltd.
TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS THE VACCINE COMPOS
TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
FILE REFERENCE: IEC010037PCT
CURRENT APPLICATION NUMBER: US/10/381,770
CURRENT FILING DATE: 2003-07-16

```
Db 901 GACTTTGCGCTCGAATGTAGTTCCGCAACCTCACCCCGGTATATACCAACACGCGGTC 960
QY 961 TCCGTTTATTCAGACACTGCTCGCACCGCTTCGTGCGGTCGCGACGGGACTGCGGAG 1020
Db 961 TCCGTTTATTCAGACACTGCTCGCACCGCTTCGTGCGGTCGCGACGGGACTGCGGAG 1020
QY 1021 CTCACACACGCGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
Db 1021 CTTACACACGCGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080
QY 1081 GTCCGTGAGATCGCGCGCGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1140
Db 1081 GTCCGTGAGATCGCGCGCGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1140
QY 1141 GCGCGCTCCGACAGAAATGATTTGCTGCGGTGGTGGCAGCTGTTCTACTCCGTCGC 1200
Db 1141 GCGCGCTTACCGACAGAAATGATTTGCTGCGGTGGTGGCAGCTGTTCTACTCCGTCGC 1200
QY 1201 GTTGCTCAGCAATGCGGAGCGGACTGTTAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Db 1201 GTTGCTCAGCAATGCGGAGCGGACTGTTAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
QY 1261 CAGGATTAAGGATTAATGCAATCCGATGACATGACCTCGGAAATCTCTGTGTTGTTAT 1320
Db 1261 CAGGATTAAGGATTAATGCAATCCGATGACATGACCTCGGAAATCTCTGTGTTGTTAT 1320
QY 1321 CAGGATTAATGATTAACCAACATGAACATGCGGCGGAGCCCTTCTCCAGCCCATCGCGC 1380
Db 1321 CAGGATTAATGATTAACCAACATGAGCAGGACCGACGACCTTCCAGCCCATCGCGC 1380
QY 1381 CTTTCTCTGCTCTGAGCTAATGATGCTGTTGGCTCTCTCTCACCGCTCGCGAT 1440
Db 1381 CTTTCTCTGCTCTGAGCTAATGATGCTGTTGGCTCTCTCTCACCGCTCGCGAT 1440
QY 1441 GACAGTCCACTTATGCTCTTCGACTGCGGCGGATTTATGTTCTGACTCTGTGACCTTG 1500
Db 1441 GACAGTCCACTTATGCTCTTCGACTGCGGCGGATTTATGTTCTGACTCTGTGACCTTG 1500
QY 1501 GTTAATGTTGGACCGCGCGGAGCCGTTGCGCGTCTGCTCGATGGAACCAAGTCA 1560
Db 1501 GTTAATGTTGGACCGCGCGGAGCCGTTGCGCGTCTGCTCGATGGAACCAAGTCA 1560
QY 1561 CTTGAGGTCGCGGCTCTCCACCATGAGAGTACTCGAAGACCTTCTTTGCTGCGG 1620
Db 1561 CTTGAGGTCGCGGCTCTCCACCATGAGAGTACTCGAAGACCTTCTTTGCTGCGG 1620
QY 1621 CTCGCGGTAAAGTCTCTTTCTGGGAGGACGACCACTAAAGCGGTACCTTATAT 1680
Db 1621 CTCGCGGTAAAGTCTCTTTCTGGGAGGACGACCACTAAAGCGGTACCTTATAT 1680
QY 1681 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATGCGCGGACCGGCTGCT 1740
Db 1681 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAAATGCGCGGACCGGCTGCT 1740
QY 1741 ATTTCCACTTACACACTAGCTGGTCTGCTGCTCCGCTCTCCATTTCTGCGGTTGCGGTT 1800
Db 1741 ATTTCCACTTACACACTAGCTGGTCTGCTGCTCCGCTCTCTATTTCCGCGGTTGCTGTT 1800
QY 1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTGAAGATACCTTGGACTACCTGCGCGGCC 1860
Db 1801 TTAGCCCCCACTCTGCGCTAGCAATGCTTGAAGATACCTTGGACTACCTGCGCGGCC 1860
QY 1861 CATATTTTATGATTTTCTGCCAGAGTGGCGGCTTGGCTTTCAGGGCTGCGCTTTC 1920
Db 1861 CATATTTTATGATTTTCTGCCAGAGTGGCGGCTTGGCTTTCAGGGCTGCGCTTTC 1920
QY 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAATACTCGGAGTTG 1980
Db 1921 CAGTCTACTGCTGAGCTTTCAGCGCTTAAAGATGAAGTGGGTAATACTCGGAGTTA 1980
QY 1981 TAG 1983
|||
```

```
Db 1980 TAG 1982
RESULT 7
US-09-769-066-3
; Sequence 3, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; McAttee, C. Patrick
; Yarbrough, Patrice O.
; Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/769,066
; FILING DATE: 24-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/542,634
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1647 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E virus (Burma) r62kDa,
; FIGURE 2
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-769-066-3
```

```
Query Match 83.1%; Score 1647; DB 9; Length 1647;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GCGGTCTCGGCGCCATGACACCCGCGCAGTGCCTGATGCTGACTCCCGGGCGGCATC 393
Db 1 GCGGTCTCGGCGCCATGACACCCGCGCAGTGCCTGATGCTGACTCCCGGGCGGCATC 60
QY 394 TTGCGCGGCGAGTATTAACCTTATCAACATCTCCCTTTACCTTTCCGTTGGCCACCGGCACT 453
Db 61 TTGCGCGGCGAGTATTAACCTTATCAACATCTCCCTTTACCTTTCCGTTGGCCACCGGCACT 120
QY 454 AACCTGGTCTTTTATGCGGCGCCCTCTAGTCCGCTTTTACCCCTTCAGGACGCAACAAT 513
Db 121 AACCTGGTCTTTTATGCGGCGCCCTCTAGTCCGCTTTTACCCCTTCAGGACGCAACAAT 180
QY 514 ACCCATATATGCGCACGGAAGCTTCTAATTATGCCCCAGTACCGGTTGCCCCGTGCCACA 573
Db 181 ACCCATATATGCGCACGGAAGCTTCTAATTATGCCCCAGTACCGGTTGCCCCGTGCCACA 240
```


QY 1694 ACAAATAAGCGGGTACCTTATAATATTAACACCACTGCTAGCGAACAACCTGCTGTC 1713
Db 1321 ACAAATAAGCGGGTACCTTATAATATTAACACCACTGCTAGCGAACAACCTGCTGTC 1380
QY 1714 GAGAAATCGCGCGGACACCGGGTCTATTTCCACTTACACCACTAGCTGGTGGTCTGCT 1773
Db 1381 GAGAAATCGCGCGGACACCGGGTCTATTTCCACTTACACCACTAGCTGGTGGTCTGCT 1440
QY 1774 CCCGCTCCATTTCTGCGGTTGCGGTTTATAGCCGCCCACTCTGCGCTAGCATTTGCTTGA 1833
Db 1441 CCCGCTCCATTTCTGCGGTTGCGGTTTATAGCCGCCCACTCTGCGCTAGCATTTGCTTGA 1500
QY 1834 GATACCTTGGACTACCTCGCGCGGCCCATACCTTTTGATGATTTCTGCCAGAGTCCGC 1893
Db 1501 GATACCTTGGACTACCTCGCGCGGCCCATACCTTTTGATGATTTCTGCCAGAGTCCGC 1560
QY 1894 CCCCTGGCTTACAGGCTGCGCTTCCAGTCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1953
Db 1561 CCCCTGGCTTACAGGCTGCGCTTCCAGTCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1954 ATGAAGGTGGTAAATACTCGGAGTTG 1980
Db 1621 ATGAAGGTGGTAAATACTCGGAGTTG 1647

RESULT 8

US-09-769-066-2
; Sequence 2, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; McAtee, C. Patrick
; Yarbough, Patrice O.
; Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/769,066
; FILING DATE: 24-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/542,634
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2058 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: Hepatitis E Virus (Mexico Strain)
; ORF-2 region
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

QY 574 ATCCGTTACCGCGGCTGCTCCCAATGCTGTCGCGGTTAGCGCAATCTCCATCTCATTC 633
Db 241 ATCCGTTACCGCGGCTGCTCCCAATGCTGTCGCGGTTAGCGCAATCTCCATCTCATTC 300
QY 634 TGGCCACAGACACACACCGCGGAGCTCGTTGATATGAATTCATTAACCTCGACGGAT 693
Db 301 TGGCCACAGACACACACCGCGGAGCTCGTTGATATGAATTCATTAACCTCGACGGAT 360
QY 694 GTTCGATTTTATGTCACCGCGGATAGCTCTGAGCTTGTGATCCCAAGTGAAGCGCTTA 753
Db 361 GTTCGATTTTATGTCACCGCGGATAGCTCTGAGCTTGTGATCCCAAGTGAAGCGCTTA 420
QY 754 CACTATCGTAACCAAGGCTGGCGCTCCGTCGAGACCTCTGCGGGTGGCTGAGAGAGGCT 813
Db 421 CACTATCGTAACCAAGGCTGGCGCTCCGTCGAGACCTCTGCGGGTGGCTGAGAGAGGCT 480
QY 814 ACTCTGGTCTTGTATGCTTGCATACATGCTCACTCGTAAATTCCTATATACTATAACA 873
Db 481 ACTCTGGTCTTGTATGCTTGCATACATGCTCACTCGTAAATTCCTATATACTATAACA 540
QY 874 CCTATACCGGTGCGCTCCGCGGCTGTTGAGCTTTGCGCTTGGAGTTGAGTTTCGCAACCTT 933
Db 541 CCTATACCGGTGCGCTCCGCGGCTGTTGAGCTTTGCGCTTGGAGTTGAGTTTCGCAACCTT 600
QY 934 ACCCCCGGTAACCAATAGCGGGTCTCCGTTATTCAGACACTGCTCGCCACCGCCTT 993
Db 601 ACCCCCGGTAACCAATAGCGGGTCTCCGTTATTCAGACACTGCTCGCCACCGCCTT 660
QY 994 CGTCCGGTTCGAGACGGGACTCGGAGCTCACACACCGCTGCTACCGCTTTATGAAG 1053
Db 661 CGTCCGGTTCGAGACGGGACTCGGAGCTCACACACCGCTGCTACCGCTTTATGAAG 720
QY 1054 GACCTCTATTTTACTAGTACTAATGCTGCTGAGATCGCGCGGATGAGCCCTCAAC 1113
Db 721 GACCTCTATTTTACTAGTACTAATGCTGCTGAGATCGCGCGGATGAGCCCTCAAC 780
QY 1114 CTGTTCAACCTTGTGACACTCTGCTGGCGGCTGCGGACAGAAATGATTTGCTGGCT 1173
Db 781 CTGTTCAACCTTGTGACACTCTGCTGGCGGCTGCGGACAGAAATGATTTGCTGGCT 840
QY 1174 GGTGGCAGCTGTTCTACTCCGTCGCTGCTCAGCCAAATGCGAGCGGACTGTTAAG 1233
Db 841 GGTGGCAGCTGTTCTACTCCGTCGCTGCTCAGCCAAATGCGAGCGGACTGTTAAG 900
QY 1234 TTGTATACATCTGTAGAGATGCTCAGCAGGATAAGGTTATGCAATCCCGCATGACAT 1293
Db 901 TTGTATACATCTGTAGAGATGCTCAGCAGGATAAGGTTATGCAATCCCGCATGACAT 960
QY 1294 GACCTCGAGAAATCTCGTGGTGTATTCAGGATTTATGAACCAACATGACAGATCGG 1353
Db 961 GACCTCGAGAAATCTCGTGGTGTATTCAGGATTTATGAACCAACATGACAGATCGG 1020
QY 1354 CCGAGCGCTTCTCCAGCCCATCGGCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1413
Db 1021 CCGAGCGCTTCTCCAGCCCATCGGCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1414 TGGCTCTCTCTCACCGCTGCGAGTATGACAGCTTATGCTTATGCTTCTGAGCTGGCCCA 1473
Db 1081 TGGCTCTCTCTCACCGCTGCGAGTATGACAGCTTATGCTTATGCTTCTGAGCTGGCCCA 1140
QY 1474 GTTATGTTTCTGACTCTGACCTTGTGTTAATGTTGCGCGGCGCGGCGGCGGCTTGGC 1533
Db 1141 GTTATGTTTCTGACTCTGTGACCTTGTGTTAATGTTGCGGCGGCGGCGGCGGCTTGGC 1200
QY 1534 CGGTGCTCGATTGACCAAGGTACACTGACGCTCGCCCTCTCCACCATCCAGCAG 1593
Db 1201 CGGTGCTCGATTGACCAAGGTACACTGACGCTCGCCCTCTCCACCATCCAGCAG 1260
QY 1594 TACTCGAAGACCTTCTTTGCTGCGCTTCGCGGTAAGCTCTCTTTCTGGAGGCGGCG 1653
Db 1261 TACTCGAAGACCTTCTTTGCTGCGCTTCGCGGTAAGCTCTCTTTCTGGAGGCGGCG 1320


```

; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Bradley, Daniel W.
; Twu, Jr-Shin
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; FILING DATE: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2055 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 7
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-165-868-4

      Query Match          69.0%; Score 1368.6; DB 15; Length 2055;
      Best Local Similarity 81.0%; Pred No. 0;
      Matches 1606; Conservative 0; Mismatches 374; Indels 3; Gaps 1

QY   1  ATGCGCCCTCGGCCTATTTTGTGTGTCCTCCTCATGTTTTTGCCCTATGCTGCCCGGCCA 60
    |||
Db   1  ATGGCCCCTAGCCCTCTTTTGTGTGTTTCTCTTGTGTTTCTGCTATGTTGTCCTCGCCA 60

QY   61  CGCCCGGTACAGCGCTCTGGCGCGCTGTGGCGCGCGCACGCGCGGTTCCTCGCGCGGTGT 120
    |||
Db   61  CGACCGGTACAGCGCTGTGGCGCGCTGTGGCGCGCGCACGCGCGGTACCGCGGTGT 120

```


346	CGACCTGCCACACCCGGGGCTGGGCGCTGACGGCTGTGGCGCCTGGCCCATGACACCTCA	405
Qy		
361	CCAGTGCTGATGTGCAGTCCCGCGCGCGCATCTTTGGCGCGGAGTATAAAGCTATCAACA	420
Db		
406	CCCGTCCGGACGGTTGATTTCTCGCGTGAATTTCTAGCGCGCAGTATAATTTGTCTACT	465
Qy		
421	TCTCCCTTACCTCTTCCGTTGCCACCGGACCTAAACCTGGTTCTTTATGCGCGCCCTCTT	480
Db		
466	TCACCCCTGACATCCTCTGTGGCTCTGCGACATAATTTAGTCTCTGTATGACGCGCCCTT	525
Qy		
481	AGTCCGCTTTTACCCCTTTAGGACGGGACCAATACCCATATAATGGCCACGGAGCTTCT	540
Db		
526	AATCCGCTCTGGCGCTGACGACGGTACTAATACTACTCAATATAGCCACAGAGGCCCTCC	585
Qy		
541	AATATACCCAGTACCGGGTTGCCGTGCCACAAATCCGTTACCGCCCGCTGGTCCCCAAAT	600
Db		
586	AATATGACAGTACCGGGTTGCCGGCTACTATCCGTTACCGTTACCGGCCCTAGTGCCTAAT	645
Qy		
601	GCTGTCCGCGGTTACGCCATCTCCATCTCAATCTGGCCAAGACCAACACCCGAGG	660
Db		
646	GCAGTTGGAGGCTATGCTATATCCATTTCTTTCTGGCCTCAAAACACCAACCCGCTACA	705
Qy		
661	TCCGTTGANTAGAAATTCOAATAAAGTCCGAGGATGTTGTAATTTAGTCAGAGCCGCGATA	720
Db		
706	TCTGTTGACATGAATTCATTTACTTCCACTGATGTGAGGATCTCTGTTCAACCTGGCATA	765
Qy		
721	GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGCCTGGGCGTCC	780
Db		
766	GCATCTGAATTTGGTCAATCCCAAGCGAGCGCTTCACTACCGCAATCAAAGTTGGGCGCTG	825
Qy		
781	GTCGAGAGCCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTGTTATGCTTTGTCATA	840
Db		
826	GTTGAGACATCTGGTGTGTTGTGAGGAGGAGCAACCTCCGCTCTTGTCATGTTATGTCATA	885
Qy		
841	CATGGCTCACTGTAATAATTCCTACTAATAACCCCTATACCGTTCGCTCCGGCTGTGG	900
Db		
886	CATGGCTCTCCAGATTAATCTCCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG	945
Qy		
901	GACTTTGGCCTTGAGCTTTGAGTTTTCGCAACTTACCCCGGTAAACCAATAACCGGGTTC	960
Db		
946	GACTTTGGCCTTAGAGCTTTGAGTTTTCGCAATCTCACACCTGTAAACCAATAACAGTGTG	1005
Qy		
961	TCCCGTTATTCAGACACTGCTCGCCACCGCCTTCGTCGGGTGGGACGGGATGCGGAG	1020
Db		
1006	TCCCGTTATTCAGACACTGCTCGTCACTCC--GCCCGAGGGCGGACGGGATGCGGAG	1062
Qy		
1021	CTCACACCAACGGCTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGCT	1080
Db		
1063	CTGACCAAACTGCAGCACCAAGGTTTCATGAAGAATCTCCACTTTACCGGCCCTTAATGGG	1122
Qy		
1081	GTCCGTGAGATCGGCGGGGATAGCCCTCAACCTGTTTCAACCTTGCTGACACTCTGCTT	1140
Db		
1123	GTAGGTGAAGTGGCGCGGGATAGCTCTAAATTAATTTAACCTTGTGACACGCTCCCTC	1182
Qy		
1141	GGCGGCTCTCCACAGAAATTTGATTTTGTGCGCTGGTGGCCAGCTGTTTCTACTCCCGTCCC	1200
Db		
1183	GGCGGCTCTCCACAGAAATTTTTCGTGCGCTGGCGGCAACTGTTTATTTCCCGCCCG	1242
Qy		
1201	GTGTCTAGCCAAATGGGAGCGGACTGTTAAGTTGTATATCATCTGTAGAGATGCTCAG	1260
Db		
1243	GTGTCTAGCCAAATGGGAGCCAAACCGTGAAGCTCTATACATCAGTGGGAGATGCTCAG	1302
Qy		
1261	CAGGATAAGGTTATGCAATCCGCATACATTTGACTCGGAGAAATCTCTGTGTGGTTATT	1320
Db		
1303	CAGGATAAGGGTGTGCTATCCCCACGATATCGATCTTGTGTGATTCGCTGTGGTCATT	1362
Qy		
1321	CAGGATTATGATAAACCAACATGAAACAGATCGCGGACGCTTCTTCAGGCCCAATCGCGC	1380
Db		
1363	CAGGATTATGAAACCAACAGATGAGGAGGATCGGCCACCCCGTCCGCTTGCCCATCTTCGG	1422
Qy		
1381	CTTTTCTCTGCTTCAGCTAATGATGTGCTTTGGCTCTCTCTCAACGCTGCGGAGTAT	1440
Db		
1423	CTTTTCTCTGCTTCGAGCAAAATGATGTACTTTGGCTGTCCCTCACTGTGCCGAGTAT	1482

RESULT 11

RESULT II
IIS-09-851-470-10

US-09-851-410-10 Application US/09851410

; Sequence 10, Application US/098

Publication No. US20030030000

; GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R

Yarbough, Patrice O

Bradley, Daniel W

Pracujący, bankier, Krawczyński.

Krawczyński
 Tam Albert

Tam, Albert
Eusebio, Eusebio

Fry, Kirk E

;
TITLE OF INVENTION: DNA

No.

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

CORRESPONDENCE
ADDRESSEE

ADDRESS
STREET. 31

; STREET: 350
CITY: PALO ALTO

; CITY: Palo Alto, CA

; STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC
MEDIUM LIFE: floppy

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; COMPILER: Datcom Pascal #1.0 Version #1.25

```

/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: US 07/681,078
/ FILING DATE: 05-APR-1991
/ APPLICATION NUMBER: US 07/505,888
/ FILING DATE: 05-APR-1990
/ APPLICATION NUMBER: US 07/420,921
/ FILING DATE: 13-OCT-1989
/ APPLICATION NUMBER: US 07/367,486
/ FILING DATE: 16-JUN-1989
/ APPLICATION NUMBER: US 07/336,672
/ FILING DATE: 11-APR-1989
/ APPLICATION NUMBER: US 07/208,997
/ FILING DATE: 17-JUN-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Petithory, Joanne R.
/ REGISTRATION NUMBER: 42,995
/ REFERENCE/DOCKET NUMBER: 4600-0183.24
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 324-0880
/ TELEFAX: (650) 324-0860
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7171 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: Composite Mexico strain
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-410-10

Query Match 69.0%; Score 1368.6; DB 10; Length 7171;
Best Local Similarity 81.0%; Pred No. 0;
Matches 1606; Conservative 0; Mismatches 374; Indels 3; Gaps 1;

QY 1 ATGCGCCCTCGGCCCTATTTGTTGCTGCTCTCTCAATGTTTTCCTATGTCGCCCGGCCA 60
Db 5117 ATGCGCCCTAGGCCCTCTTTGCTGTTGTTCTCTCTTTGCTCTATGTTGCGCGGCCA 5176

QY 61 CGGCCGCTGACCGCTCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 120
Db 5177 CGGACCGCTGACCGCTCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGT 5236

QY 121 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 180
Db 5237 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 5296

QY 181 CCCTTCGCCCGCCGATGTACCCGCTGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCC 240
Db 5297 CCCTTTGCGCCGAGAGTTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCC 5356

QY 241 CGACCACTCGGCTCGGCTGGGGTGACAGCGCCGCGCGCGCGCGCGCGCGCGCGCGT 300
Db 5357 CGGCGACTTGGCTCACTTGGCGGATCAGGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGCGT 5416

QY 301 AGACCTACCAAGCTGGGGCGCGCGCTGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 360
Db 5417 CGACTGCGACAGCGGGGCTGGCGCGCTGACGGCTGCGCGCGCGCGCGCGCGCGCGCGT 5476

QY 361 CCAGTGCTGATGCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 420
Db 5477 CCGGTCGCGGAGTGAATCTCGCGGTGCAATTCAGCGCGCGCGCGCGCGCGCGCGCGT 5536

QY 421 TCTCCCTTACTCTCTCGTGGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 480
Db 5537 TCACCCCTGACATCTCTGTGGCGCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGT 5596

QY 481 AGTCGCGCTTTTACCCCTTCAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 540
Db 5597 AATCCGCGCTCTGCGCTGCGAGGCGGTAATACTACATTTATGGCGCGCGCGCGCGT 5656

QY 541 AATTATGCCAGTACCGGGTTGCCGTGCCAATCGTTACCGCGCGCTGGTCCCAAT 600
Db 5657 AATTATGCCAGTACCGGGTTGCCGTGCCAATCGTTACCGCGCGCTGGTCCCAAT 5716

QY 601 GCTGTCGGGGTTACGCCATCTCTCATCTCTGCGCCACAGACCAACCAACCGCGAGC 660
Db 5717 GCAGTTGGAGGCTATGCTATATCAATTTCTTCTGCGCTCAACACCAACCAACCGCGAGC 5776

QY 661 TCGTTGATATGAATTAATTAACCTCGAGGATGTTGATTTTATGTCAGCGCGCGAT 720
Db 5777 TCTGTTGACATGAATTTCAATTTACTTCCACTGATGTGAGGATTTCTGTTCAACCTGCGATA 5836

QY 721 GCCTCTGAGCTTGTGATCCCAAGTACGCGCTACACTATCGTAACCAAGGCTGGCGTCC 780
Db 5837 GCATCTGAATTTGCTCATCCAAAGCGAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 5896

QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTTCATA 840
Db 5897 GTTGAGACATCTGGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5956

QY 841 CATGCTCACTCGTAAATTCCTATATCAATACACCTATACCGGTGCGCTCGGGCTGTTG 900
Db 5957 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 6016

QY 901 GACTTTGGCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGTC 960
Db 6017 GACTTTGGCCTTGAGCTTGAGTTTCGCAATCTCACCACTCTGTAACCAATACAGGTG 6076

QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTCGCGGTGCGGAGCGGAGTGGCGAG 1020
Db 6077 TCCCGTTATTCAGCACTGCTCGCTCACTCC---GCCCGAGGGGCGGAGTGGCGAG 6133

QY 1021 CTCACACACCGGCTGCTACCGCTTTATGAAGGAGCTTATTTACTAGTACTAATGT 1080
Db 6134 CTGACCAACACTGCGAGCGGAGGTTTATGAAGATCTCCACTTTACCGGCGCTTAAATGG 6193

QY 1081 GTCGCTGAGATCGGCGCGGAGTAGCCCTCAACCTGTTCAACCTTCTGCTGAGACTCTGTT 1140
Db 6194 GTAGTGAAGTCGGCGCGGAGTAGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 6253

QY 1141 GCGCGCTCGCGGACAGAAATGATTTGCTGGGTGGCGGAGCTGTTCTTCTTCCCGTCCC 1200
Db 6254 GCGCGCTCGCGGACAGAAATTAATTTGCTGGGTGGCGGAGCTGTTTATTTCCCGCGCG 6313

QY 1201 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGTATATCATCTGTAGAGAGTCTCAG 1260
Db 6314 GTTGTCTCAGCAATGCGGAGCGGAGCTGTTAAGTTGTATATCATCTGTAGAGAGTCTCAG 6373

QY 1261 CAGGATAGGGTATTTGCAATCCCGCATGACATTCAGCTCGGAGATCTCGTGTGGTTAT 1320
Db 6374 CAGGATAGGGTATTTGCAATCCCGCATGACATTCAGCTCGGAGATCTCGTGTGGTTAT 6433

QY 1321 CAGGATATGATAACCAACATGAACAAGATCGGCGGAGCTTCTTCCAGCGCGGAGTCCGCG 1380
Db 6434 CAGGATATGATAACCAACATGAACAAGATCGGCGGAGCTTCTTCCAGCGCGGAGTCCGCG 6493

QY 1381 CCTTCTCTGCTCTCGAGCTTAATGATGCTGTTGGTCTCTCTCAACGCTGCGCGAGTAT 1440
Db 6494 CCTTCTCTGCTCTCGAGCTTAATGATGCTGTTGGTCTCTCTCAACGCTGCGCGAGTAT 6553

QY 1441 GACCACTCCACTTATGCTCTCGAGCTGGCGGAGTATGTTTATGTTCTGACTCTGTGACCTTG 1500
Db 6554 GACCACTCCACTTATGCTCTCGAGCTGGCGGAGTATGTTTATGTTCTGACTCTGTGACCTTG 6613

QY 1501 GTTAATGTTGCGAGCGGCGGAGCGGCTTGGCGGCTGCTGCGATTCGATTCGACCAAGGTCACA 1560
Db 6614 GTGAATGTTGCGAGCTGGCGGAGCGGCTGAGCGGAGTCTGCTGACTGCTCAAGTCAAC 6673

QY 1561 CTTGACCGGTGCGCGCGCTCTCAACATCCAGAGTACTCGAAGACCTTCTTTGCTCGCG 1620
Db 6674 CTTGACCGGTGCGCGCGCTCTCAACATCCAGAGTACTCGAAGACCTTCTTTGCTCGCG 6733

QY 1621 CTCCGCGTGAAGCTCTCTTCTGCGGAGGAGGAGCAAACTAAAGCGGAGTACCTTATAAT 1680

Db	6734	CTTCGTGGCAGAGCTCTCCCTTTTGGGAGGCGGCACAAACAAAGCAGGTTATCCCTTATAAT	6793	QY	301	AGACCTACACAGCTGGGCGCGCGCTAAACCGGGTGGCTCCCGGCCCATGACACCCCG	360
QY	1681	TATAACACACAGCTAGGACCAACTGCTTGTGAGAAATGCCCGGGCACCGGGTCGCT	1740	Db	5417	CGACCTGCCACAGCGCGGGCTGGCGGCTGACGGCTGTGGCGCTGCCCATGACACCTCA	5476
Db	6794	TATAATACCTACTGCTAGTGACAGATCTGATTGAAATGCTGCGGCCATCGGGTCGCC	6853	QY	361	CCAGTGCCTGATGTGCACTCCCGCGCGCACTCTTGGCGCGGCACTAATCACTATCAACA	420
QY	1741	ATTTCCACTTACACCACTAGCCTGGGTGCTGCTCCCGTCTCAATTTCTGCGGTTCGCGTT	1800	Db	5477	CCGCTCCCGGAGGTTGATTCTCGCGGTGCAATTTCTACGCGCGCAGTATAATTTGTCTACT	5536
Db	6854	ATTTCAACCTATACCACAGGCTTGGGCGCGTCCGGTCGCCATTTCTGCGCGCGGTT	6913	QY	421	TCCTCCCTTACCTCTTCCGTGGCCACCGGCACTAACCCTGGTTCTTTATCGCGCCCTCTT	480
QY	1801	TTAGCCCCCACCCTGCGCTAGCATCTTGGAGTACCTTGGACTACCTCCCGCGCC	1860	Db	5537	TCACCCCTGACATCCTCTGTGGCTCTGGCACTAATTTAGTCTCTGTATGACGCCCTCTT	5596
Db	6914	TTGGCTCCACGCTCCGCCCTGCTCTGTGGAGGATCTTTTGAATATCCGGGCGGGCG	6973	QY	481	AGTCGCTTTTACCCCTTCAGGACGGCACCAATACCATATATATGSCCAGGAAGCTTCT	540
QY	1861	CATACCTTTTGATGATTTCTGCCAGAGTGGCGGCCCTTGGCCCTTCAGGGCTGCGCTTTC	1920	Db	5597	AATTCGCTCTCGCGCTGCAGACGGTACTATCTACATTTATGSCCAGAGGCTCC	5656
Db	6974	CACACATTTGATGACTTCTGCCCTGATGCGCGGCTTTAGGGCTCCAGGGTTGCTTTC	7033	QY	541	AATATGCCCACTACCGGGTCCCGGTGCACAATTCGGTTACGCCCGCTGGTCCCAAT	600
QY	1921	CAGTCTACTGCTCGCTGAGCTTCAGCGCCCTTAAGATGAAGTGGGTAAACTCGGGAGTTG	1980	Db	5657	AATATGCACAGTACCGGGTTCGGCGCTACTATCCGTTACCGGCGCTTAGTGCCTAAT	5716
Db	7034	CAGTCAACTGTGCTGAGCTCAGCGCCCTTAAAGTTAAGGTGGGTAAAACTCGGGAGTTG	7093	QY	601	GCTGTGCGGGTTACGCCATCTCCATCTCATTTGSCCAGACACCAACCCCGGACG	660
QY	1981	TAG 1983		Db	5717	GCAAGTGGAGGCTATGCTATATCCATTTCTTTCTGGCCCTCAACACACACACCCCTACA	5776
Db	7094	TAG 7096		QY	661	TCCGTTGATATGAATTCAAATAACCTCGACGATGTTCTGTTATTTAGTCCAGCCCGGATA	720
RESULT 12	US-10-239-090A-49			Db	5777	TCTGTTGACATGAATTCCTACTTCCACTGATGTCAGGATTTCTTTCAACCTGGCATA	5836
	Sequence 49, Application US/10239090A			QY	721	GCTCTGAGCTTGTGATCCCAAGTGAAGCGCTACATATGTTAACCAAGCTGGCGCTCC	780
	Publication No. US20040101820A1			Db	5837	GCATCTGAATTTGGTCAATCCAAAGCAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG	5896
	GENERAL INFORMATION:			QY	781	GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGCTTCTGTTATGCTTTGCATA	840
	APPLICANT: KAHUSHIKI KALSHA TOSHIBA			Db	5897	GTTGAGACATCTGTTGTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	5956
	TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co			QY	841	CATGGCTCACTCGTAAATTCCTATATAATACACCTTATACCGGTGCGCTTCGGGTGTTG	900
	TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,			Db	5957	CATGGCTCTCCAGTTAACTTCTTATACCAATACCCCTTATACCGGTGCGCTTCGGCTACTG	6016
	TITLE OF INVENTION: detecting hepatitis E virus using the same			QY	901	GACTTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT	960
	FILE REFERENCE: 02S0741P			Db	6017	GACTTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT	6076
	CURRENT APPLICATION NUMBER: US/10/239, 090A			QY	961	TCCCGTTATTTCCAGCACTGCTCGCCACCGCTTCTGTCGCGTGGCGGAGGAGGAGGAGGAG	1020
	PRIOR FILING DATE: 2003-07-24			Db	6077	TCCCGTTATTTCCAGCACTGCTCGCCACCGCTTCTGTCGCGTGGCGGAGGAGGAGGAGGAG	6133
	PRIOR APPLICATION NUMBER: JP2001-191837			QY	1021	CTCACCACCAACGGCTGTCTACCCGCTTTATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1080
	PRIOR FILING DATE: 2001-06-25			Db	6134	CTGACCACTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	6193
	NUMBER OF SEQ ID NOS: 57			QY	1081	GTCGGTGAGATCGGCGCGGGATAGCCCTCACCTGTTCAACCTGTTCAACCTGTTCAACCTGTT	1140
	SEQ ID NO 49			Db	6194	GTAAGTGAAGTGGCGCGGGATAGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA	6253
	LENGTH: 7180			QY	1141	GGCGGCTCGCGCAGAGAAATGATTTCGTGGCTGGTGGCGGAGGAGGAGGAGGAGGAGGAGG	1200
	TYPE: DNA			Db	6254	GGCGGCTCGCGCAGAGAAATGATTTCGTGGCTGGTGGCGGAGGAGGAGGAGGAGGAGGAGG	6313
	ORGANISM: Hepatitis E Virus Mexico (M74506)			QY	1201	GTTGTCTCAGCCAAATGGCGAGCGGAGCTGTTAAGTTGTTATATCATCTGTAGAGAAATGCTCAG	1260
	US-10-239-090A-49			Db	6314	GTTGTCTCAGCCAAATGGCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	6373
Query Match	69.08; Score 1368.6; DB 17; Length 7180;			QY	1261	CAGGATGAAGGTTATGCAATCCCGCATGACATGACCTCGGAGAACTCTCGTGTGGTTATT	1320
Best Local Similarity	81.0%; Pred. No. 0;			Db	6374	CAGGATGAAGGTTATGCAATCCCGCATGACATGACCTCGGAGAACTCTCGTGTGGTTATT	6433
Matches 1606; Conservative	0; Mismatches 374; Indels 3; Gaps 1;			QY	1321	CAGGATTTATGATTAACCAACATGAACAAAGATCGGCGGAGCGGCTTCTCCAGCCCAATCGCGC	1380
QY	1	ATCGGCTCGGCTATTTGTTGCTGCTCTCTCATGTTTTCCTATGCTTTCCTATGCTGCGCGGCCA	60	Db	6434	CAGGATTTATGATTAACCAACATGAACAAAGATCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG	6493
Db	5117	ATCGGCTCGGCTATTTGTTGCTGCTCTCTCATGTTTTCCTATGCTTTCCTATGCTTTCCTATG	5176	QY	1381	CCTTTCTCTGTCTTCGAGCTAAATGATGTGCTTTGGTCTCTCTCTCACCAGCTGCCAGTAT	1440
QY	61	CCGCGCGGCTCAGCGCTGCGCGCGCTGCTGCGCGCGGAGCGAGCGCGGTTCCGGCGGTGGT	120				
Db	5177	CCGACCGGCTCAGCGCTGCGCGCGCTGCTGCGCGCGGAGCGAGCGCGGTTACCGCGGTGGT	5236				
QY	121	TTCTGGGGTGACCGGGTTGATTTCTAGCCCTTTCGCAATCCCTATATTTCAACCAACCAAC	180				
Db	5237	TTCTGGGGTGACCGGGTTGATTTCTAGCCCTTTCGCAATCCCTATATTTCAACCAACCAAC	5296				
QY	181	CCCTTTCGCGCCCGAGTTCACCGCTGCGCGCGGGGCTGAGACCTCGTGTTCGCAACCCGCGC	240				
Db	5297	CCCTTTCGCGCCCGAGCTTTCGCGCTGCGCTGCGGTCTGGACCTCGCTTTCGCAACCCGCGC	5356				
QY	241	CGACCACTCGGCTCGGCTTGGCGGTGAACAGCGCCAGCGCCCGCGGTTCGCTCAGCTGCT	300				
Db	5357	CGGCACTTGGCTCCACTTGGCGAGATCAGGCGCCAGCGCCCTCGGCTGCTCCCGTGGC	5416				

Db	6313	GGTGGTCTGCCGACAGAAATGATTTCTGTGGCTGGGGCCAGTGTGTTCTACTCCGCCCT	6372
Qy	1201	GTTGTCTCAGCCAAATGGCGAGCCAGCATGTTAAAGTTGTATATACATCTGTAGAGATGCTCAG	1260
Db	6373	GTCGTCTCAGCCAAATGGCGAACCAGCATGTAAAGTTATATACATCTGTGTGAGAATGACACAG	6432
Qy	1261	CAGGATAAGGGTATTGCAATCCGCAATGACCTCGAGAAATCTCGTGTGGTTATT	1320
Db	6433	CAGGATAAGGGTATCACATCCACAGATATAGACCTCGGTGATTTCTGTGTGGTCAATC	6492
Qy	1321	CAGGATTATGATTAACCAACATGAACAAAGATCGGCCGACGCTTCTCGAGCCCCCATCGCGC	1380
Db	6493	CAGGACTATGATAACCAAGCAGCAGCAGGACCGACCTACCCGCCCTCACCGGCCCTCTCGT	6552
Qy	1381	CTTTTCTCTGTCTTCGAGCTAAATGATGTGCTTTGGCTCTCTCTCAACGCTGCCGAGTAT	1440
Db	6553	CCATTTCTGGTCTTTTCGCGCTAAATGATGTTCTGTGGCTCTCTCTTACCGCTGCCGAGTAT	6612
Qy	1441	GACCAGTCCACTTATGGCTCTTCGACTGGCCCCAGTTTATGTTTCTGACTCTGTGAACCTTG	1500
Db	6613	GACCAGACTAGTATGGGTCACTTACCAACCCCATGTATGTCTCGGACACAGTCACATTT	6672
Qy	1501	GTTAATGTTGGACCGGCGCGCAGGCGGTTGCCCGGTGCGTTCGATTTGGACCAAGGTACACA	1560
Db	6673	GTTAAACGTGGCCACTGGTGTCTCAGGCTGTTGCTCGCTCCCTTGATTTGGTCTAAAGTTACC	6732
Qy	1561	CTTGACCGTCCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTCTGCCG	1620
Db	6733	TTGACCGGTCCGCCCTCCACTACTATTCAGCAGTATTTCTAAGACATTTCTAGCTTCTCCG	6792
Qy	1621	CTCCGCGGTAAGCTCTCTTTCTGGGAGCAGGSCACAACTAAAGCCGGTACCCTTATAAT	1680
Db	6793	CTTCGTGGAAAGTTGTCITTTTGGGAGGCTGGCAGCACCAAGGCTGGTTATCTCTTATAAT	6852
Qy	1681	TATAACACACTGCTTAGCGACCAACTGCTTGTGAGATGCCCGGGCACCCGGTGCCT	1740
Db	6853	TATAATACAACTGCTAGTAGATCAGATTTTAATAGAGAACGAGCTGGCCATCTGTGTGCT	6912
Qy	1741	ATTTCCCACTTACACCACTAGCTGGGTGCTCGTCCCGTCTCCATTTCTGGGTTGCCGCTT	1800
Db	6913	ATCTCTACTTTACATACAGACTGGCGCGCGGCCACTTCTATCTCTGGGTCCGGTGTG	6972
Qy	1801	TTAGCCCCCACTCTCGCTAGCAATGTTGTGAGATACCTTTGACATCCCTGCCCGCGGCC	1860
Db	6973	TTAGCCCCGCACTCAGCCCTTGTCTTCTTGAGGATACCGCGGATACCCCGCTCTGTGCC	7032
Qy	1861	CATATCTTTTGATGATTTCTGCCAGAGTGGCGGCCCTTTGACCTTCAGGGCTGCGCTTC	1920
Db	7033	CATACTTTTGATGACTTCTGCCCGGAATGTGCGACCCCTTGGTTTGCAGGGCTGCGCTTC	7092
Qy	1921	CAGTCTACTGTGCGCTGAGCTTCAGGCGCTTAAAGATGAAGGTGGGTAAACCTCGGAGTTG	1980
Db	7093	CAGTCTACTATGTCTGAGCTTCAGCGCCTTAAATGAAGGTAGGTAAACCCGGAGTTC	7152
Qy	1981	TA	1982
Db	7153	TA	7154

RESULT 14
US-10-239-090A-47
; Sequence 47, Application US/10239090A
; Publication No. US20040101820A1
; GENERAL INFORMATION:
; APPLICANT: KABUSHIKI KAISHA TOSHIBA
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus CD
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,
; TITLE OF INVENTION: detecting hepatitis E virus using the same
; FILE REFERENCE: 02S0741P
; CURRENT APPLICATION NUMBER: US/10/239,090A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837

Db	6967	CTAGCCCCACACTCGGCCCTCGCGGCTCTAGAGGACACCGTTGATTACCCCGCCGCGCT	7026
Qy	1861	CATACTTTTGATGATTTCTGCCAGAGTGCAGCCCTTGCGCTTCAGGCTTCGCTTTC	1920
Db	7027	CACACTTTTGATGATTTTGCCTCGAGTGCCTACCCCTCGTTTTCAGGGTTTGTGCATTC	7086
Qy	1921	CAGTCTACTGTCGTGTAGCTTTCAGCGCTTTAAGATGAAGTGGGTAAAACCTCGGAGTTG	1980
Db	7087	CAGTCTACTATCGTGTAGCTTCAGCGCTTTAAATGAAGGTAGTAAAAACCGGGAGTCT	7146
Qy	1981	TA 1982	
	7147	TA 7148	
Db			

Search completed: August 21, 2004, 22:41:57
Job time : 895.671 secs

D _b	5887	GCTCCGAGTTAGTCAATCCCTAGTAGCGCCTCCACTACCGCAATCAAGGCTGGCGCTCT	5946
Q _y	781	GTCGAGACCTCTCGGGTGCTGAGGAGGCTACCTCTGGTCTCTGTGTATAGCTTTGGATA	840
D _b	5947	GTTGAGACCACGGGTGTGGCCGAGGAGGCTACTTCCGGTCTCGTAATGCTTTGCATT	6006
Q _y	841	CATGGCTCACTCGTAATTCCTATACATAACCCATACACCTATACCGGTGCCCTCGGGCTGTG	900
D _b	6007	CATGGTTCTCTGTAAATTCCTATACATAACCCATACACCTATACCTGGTGCACTGGGGTCCITT	6066
Q _y	901	GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCCGGGTAAACCAAATACCGGGTTC	960
D _b	6067	GATTTTGCAATAGAGCTTGAAATTTAGAAATCTGACACCCGGGAACAACAACCCGTGTT	6126
Q _y	961	TCCCGTTATTCACGCACTGCTCCGCCACCGCCTTCGTGGGTGGCGGACGGGACCTCCGAG	1020
D _b	6127	TCCCGGTATACGACACAGCCCGCTCACCGGCTCGGACGGGTGCTGATGGGACTGCTGAG	6186
Q _y	1021	CTCACCAACCGGCTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGTT	1080
D _b	6187	CTCACCAACCGAGCAGCACAGCTTTTATGAAGAATTTGCAATTTTACTGCGCAGAACGTT	6246
Q _y	1081	GTCGGTGAGATCGGCCCGGGGATAGCCCTCACCCCTGTTCAACCTTGCTGACACTCTGTTT	1140
D _b	6247	GTTGGTGAGTGGGTCTGGCATCGCTCGACATTTGTTAACTCGCTGATACGCTTCTC	6306
Q _y	1141	GGCGGCTTCGCGACAGAAATGATTTTGTGGCTGGTGGCCAGCTGTTCTACTCCGTCCTC	1200
D _b	6307	GGTGGTTTACCGBACAGAATTGATTTCTGTCGCGCGGGGTTCAGCTGTTTACTCCGCGCT	6366
Q _y	1201	GTTGTCTCAGCCAATGGGAGCGCACTGTTAACTTGTTATACATCTGTAGAGAATGCTCAG	1260
D _b	6367	GTTGTCTCAGCCAATGGGAGCGCAACAGTAAAGTTTATACATCTGTTGAGAACGGCGAG	6426
Q _y	1261	CAGGATAGGGTATTTGCAATCCCGCATGACATTTGACCTCGGAGAATCTCGTGTGGTTATT	1320
D _b	6427	CAAGATAGGGCAATTACATCCACAGATATAGATTTGGGTGACTCCGCTGGTGGTTATC	6486
Q _y	1321	CAGGATATGATTAACCAACATGAAACAGATCGGCGGAGGCTTCTTCGAGCCCATCGCG	1380
D _b	6487	CAGGATATGATTAACCAACAGATCGGCGCTACCCCGTCACTGCGCCCTCCCG	6546
Q _y	1381	CTTTCTCTGTCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGTGCAGAGTAT	1440
D _b	6547	CTTTCTCAGTTCTTCGTGGCAATGATGTTCTGTGGCTCTCCCTCACCGCGCTGAGTAC	6606
Q _y	1441	GACCACTCACTTATGGCTCTTCGACTGGCCCGAGTTTTATGTTTCTGACTCTGTGACCTTG	1500
D _b	6607	GACCAGACTACATATGGGTCTGTCACCAACCTATGATGTTCTCCGACAGGTCACGCTC	6666
Q _y	1501	GTTAATGTTGACCGGCGGAGCGGCTGGCCGGTGGCTCGATGTTGGACCAAGGTACA	1560
D _b	6667	GTTAATGTGGCACTGGTGTCTCAGGCTGTGCGCGCTCTCTCGAATGGTCTAAAGTTACC	6726
Q _y	1561	CTTGAGGTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACTTCTTTGTCTCTGCCG	1620
D _b	6727	TTGGATGCGCGCCCTTACTACTATTCCAGCAGTATTTCTAAGACATTTCTATGTCTCCCG	6786
Q _y	1621	CTCCGCGGTAAAGTCTCTTTTCTGGAGGACGGCAACAATAAGCCGGGTACCTTTATAAT	1680
D _b	6787	CTTCGCGGAAGTGTCTTTTGGGAGGCTGGCACTACCAAGGCGGCTACCCGTATTAAT	6846
Q _y	1681	TATAACACAATGCTAGGACCAACTGCTTGTGAGATGCGCGGGCACCGGTCGCT	1740
D _b	6847	TATAATACTACTGCTAGTGATCAAAATTTGATTGAGAACGCGCGCGCACCGGTGCTGCT	6906
Q _y	1741	ATTTCCACTTACACCACTAGCTGGGTCTGGTCCGCTCTCCATTTCTGGGTTGCCGTT	1800
D _b	6907	ATTTCTACCTATACTACTAGTTTGGGTGCGGCTCTACTCTGATCTCTGCGTCCGTT	6966
Q _y	1801	TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTTGAGACTACCTTCCCGGCC	1860

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 08:42:47 ; Search time 7630.72 Seconds
(without alignments)
11246.543 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 1980

Sequence: 1 ATGGCGCCTAGGCTCTTTT.....GTAAACTCGGAGTTCGTAG 1980

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	1980	100.0	1980	6	I73140	I73140 Sequence 2
2	1980	100.0	2055	6	AR002161	AR002161 Sequence
3	1980	100.0	2055	6	AR170405	AR170405 Sequence
4	1980	100.0	2055	6	AR232466	AR232466 Sequence
5	1980	100.0	2100	6	AR002159	AR002159 Sequence
6	1980	100.0	2100	6	AR232464	AR232464 Sequence
7	1980	100.0	2100	6	AR150916	AR150916 Sequence
8	1980	100.0	2171	6	AR207634	AR207634 Sequence
9	1980	100.0	7180	6	BD189811	BD189811 Polynucle
10	1980	100.0	7180	14	HPENSSP	M74506 Hepatitis E
11	1644	83.0	1644	6	AR170407	AR170407 Sequence
12	1378.2	69.6	1983	6	E17107	E17107 Hepatitis E
13	1378.2	69.6	7138	14	HPEA	M80581 Hepatitis E
14	1378.2	69.6	7168	6	AR139826	AR139826 Sequence
15	1378.2	69.6	7168	6	AR167470	AR167470 Sequence
16	1378.2	69.6	7168	6	AR234194	AR234194 Sequence
17	1378.2	69.6	7168	6	BD084498	BD084498 Recombina
18	1378.2	69.6	7194	6	E17109	E17109 Hepatitis E
19	1378.2	69.6	7204	14	AF444002	AF444002 Hepatitis
20	1376.6	69.5	2173	14	AF065061	AF065061 Hepatitis
21	1376.6	69.5	7204	14	AF444003	AF444003 Hepatitis
22	1376.6	69.5	7212	14	AY230202	AY230202 Hepatitis
23	1373.4	69.4	7193	14	HPEGENA	L25547 Hepatitis E
24	1373.4	69.4	7221	14	HPEORFS	L25595 Hepatitis E
25	1371.8	69.3	1984	6	AR278906	AR278906 Sequence
26	1371.8	69.3	7207	14	HPECG	D11092 Hepatitis E
27	1368.6	69.1	1983	6	I73139	I73139 Sequence 1
28	1368.6	69.1	1983	14	AF124407	AF124407 Hepatitis
29	1368.6	69.1	2049	6	AR002160	AR002160 Sequence
30	1368.6	69.1	2049	6	AR170404	AR170404 Sequence
31	1368.6	69.1	2049	6	AR232465	AR232465 Sequence
32	1368.6	69.1	2094	6	AR002158	AR002158 Sequence
33	1368.6	69.1	2094	6	AR232463	AR232463 Sequence
34	1368.6	69.1	7195	6	AR150915	AR150915 Sequence
35	1368.6	69.1	7195	6	AR207633	AR207633 Sequence
36	1368.6	69.1	7207	14	HPE5VP	M73218 Hepatitis E
37	1367	69.0	7170	14	AY204877	AY204877 Hepatitis
38	1365.4	69.0	7143	14	AF185822	AF185822 Hepatitis
39	1365.4	69.0	7194	14	HEVORFS	X99441 Hepatitis E
40	1365.4	69.0	7194	14	HPEHEV	D10330 Hepatitis E
41	1365.4	69.0	7194	14	HPEUIGH	D11093 Hepatitis E
42	1363.8	68.9	7176	14	HPEGENSA	L08816 Hepatitis E
43	1363.8	68.9	7200	14	HPECEGNOM	M94177 Hepatitis E
44	1362.2	68.8	7199	14	AF051830	AF051830 Hepatitis
45	1362.2	68.8	7202	14	HEVCG	X98292 Hepatitis E

ALIGNMENTS

RESULT 1	I73140	1980 bp	DNA	linear	PAT 03-APR-1998
LOCUS	I73140	Sequence 2 from patent US 5686239.			
DEFINITION	I73140	GI:3009279			
ACCESSION	I73140				
VERSION	I73140.1				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1980)				
AUTHORS	Reyes,G.R., Tam,A.W. and Yarbough,P.O.				
TITLE	Hepatitis E virus peptides and methods				
JOURNAL	Patent: US 5686239-A 2 11-NOV-1997;				
FEATURES	Location/Qualifiers				

Db	1861	ACATTTGATGACTTCTGCGCTGAATGCGCGCTTTAGGCTCCAGGGTTGTGCTTCCAG	1920
Qy	1921	TCAACTGTCGCTGAGCTCCAGCGCCTTAAAGTTAAAGTGGGTAAACCTCGGAGTTGTAG	1980
Db	1921	TCAACTGTCGCTGAGCTCCAGCGCCTTAAAGTTAAAGTGGGTAAACCTCGGAGTTGTAG	1980
RESULT 3			
LOCUS	AR170405	2055 bp	DNA linear PAT 17-DEC-2001
DEFINITION	Sequence 2 from patent US 6291641.		
ACCESSION	AR170405		
VERSION	AR170405.1 GI:17908364		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2055)		
AUTHORS	Fuerst,T.R., McAtee,C.Patrick., Yarbough,P.O. and Zhang,Y.-F.		
TITLE	Hepatitis E virus antigens and uses therefor		
JOURNAL	Patent: US 6291641-A 2 18-SEP-2001;		
FEATURES	Location/Qualifiers		
source	1..2055		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match 100.0%; Score 1980; DB 6; Length 2055;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGCGCCCTAGGCTCTTTGCTGTTGTTCTCTGTTGTTCTGCTATGTTGCCCGGCCA	60
Db	1	ATGCGCCCTAGGCTCTTTGCTGTTGTTCTCTGTTGTTCTGCTATGTTGCCCGGCCA	60
Qy	61	CCGACCGGTACGCGGTCTGCGCGCGCTGCTGGCGCGCGCAGCGCGGTACCGCGGTGT	120
Db	61	CCGACCGGTACGCGGTCTGCGCGCGCTGCTGGCGCGCGCAGCGCGGTACCGCGGTGT	120
Qy	121	TTCTGGGGTGACCGGGTGTATTTCTAGCCCTTCGCAATCCCTATATTCATCAACCAAC	180
Db	121	TTCTGGGGTGACCGGGTGTATTTCTAGCCCTTCGCAATCCCTATATTCATCAACCAAC	180
Qy	181	CCCTTTGCCCCAGAGCTTTCGCGTTCGCTGCGGGTCTGACCTCGCCATCACAGCC	240
Db	181	CCCTTTGCCCCAGAGCTTTCGCGTTCGCTGCGGGTCTGACCTCGCCATCACAGCC	240
Qy	241	CGGCCACTTTGGCTCCACTTTGGCGAGATCAGGCCCGCCCTCGCTGCTCCCGTGC	300
Db	241	CGGCCACTTTGGCTCCACTTTGGCGAGATCAGGCCCGCCCTCGCTGCTCCCGTGC	300
Qy	301	CGACCTGCCACAGCGGGCTGGCGCGCTGACGCGCTGGCGCTGCCATGACACCTCA	360
Db	301	CGACCTGCCACAGCGGGCTGGCGCGCTGACGCGCTGGCGCTGCCATGACACCTCA	360
Qy	361	CCCGTCCCGGAGTGTGATTTCTCGCGGTGCAATTTCTACGCGCGCAGTATAATTTGTCTACT	420
Db	361	CCCGTCCCGGAGTGTGATTTCTCGCGGTGCAATTTCTACGCGCGCAGTATAATTTGTCTACT	420
Qy	421	TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTTAGTCTCTGTATGACGCCCTTT	480
Db	421	TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTTAGTCTCTGTATGACGCCCTTT	480
Qy	481	AATCCGCTCTGCGCTGCGAGACGGTACTAATACTCACATTTATGGCCACAGAGCGCTCC	540
Db	481	AATCCGCTCTGCGCTGCGAGACGGTACTAATACTCACATTTATGGCCACAGAGCGCTCC	540
Qy	541	AATTATGACAGTACCGGGTGGCGGCTACTATCCGTTACCGGCCCTAGTGCCTAAT	600
Db	541	AATTATGACAGTACCGGGTGGCGGCTACTATCCGTTACCGGCCCTAGTGCCTAAT	600
Qy	601	GCAGTTGGAGGCTATGCTATATCCATTTCTTTCTGGGCTCAAAACCAACCAACCCCTACA	660

Db	601	GCAGTTGGAGGCTATGCTATATCCATTTCTTTCTGGCCTCAAAACCAACCCCTACA	660
Qy	661	TCGTTGATGATGAATTCATTTCTTCCATGATGTCAGGATTTCTGTTCAAACCTGGCATA	720
Db	661	TCGTTGATGATGAATTCATTTCTTCCATGATGTCAGGATTTCTGTTCAAACCTGGCATA	720
Qy	721	GCATCTGAATTTGCTCATCCCAAGCGAGCGCTTCACTACCGCAATCAAGTTGGCGCTCG	780
Db	721	GCATCTGAATTTGCTCATCCCAAGCGAGCGCTTCACTACCGCAATCAAGTTGGCGCTCG	780
Qy	781	GTTGAGACATCTGTTGTTCTGAGGAGGAAGCCACTCCGGTCTTGTCTCATGTTATGCATA	840
Db	781	GTTGAGACATCTGTTGTTCTGAGGAGGAAGCCACTCCGGTCTTGTCTCATGTTATGCATA	840
Qy	841	CATGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG	900
Db	841	CATGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG	900
Qy	901	GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAAACCAATACAGTGTG	960
Db	901	GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAAACCAATACAGTGTG	960
Qy	961	TCCCGTTACTCCAGCACTGCTCTCATCTCCGCCGAGGGGCGAGCGGAGCTCGGAGCTG	1020
Db	961	TCCCGTTACTCCAGCACTGCTCTCATCTCCGCCGAGGGGCGAGCGGAGCTCGGAGCTG	1020
Qy	1021	ACCAACTGCGAGCCACCGAGTTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA	1080
Db	1021	ACCAACTGCGAGCCACCGAGTTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA	1080
Qy	1081	GGTGAAGTCGGCGGGGATAGCTCTAACTTAACTTAACTTAACTTAACTTAACTTAACTT	1140
Db	1081	GGTGAAGTCGGCGGGGATAGCTCTAACTTAACTTAACTTAACTTAACTTAACTTAACTT	1140
Qy	1141	GGGCTCCCGACAGAAATTAATTTGCTGCGTGGCGGGCAACTGTTTATTTCCGCCCGGTT	1200
Db	1141	GGGCTCCCGACAGAAATTAATTTGCTGCGTGGCGGGCAACTGTTTATTTCCGCCCGGTT	1200
Qy	1201	GTCTCAGCCAAATGGCGGCAACCGTGAAGCTCTATACATCAGTGGAGATGCTCAGCAG	1260
Db	1201	GTCTCAGCCAAATGGCGGCAACCGTGAAGCTCTATACATCAGTGGAGATGCTCAGCAG	1260
Qy	1261	GATAAGGTTGTTCTATCCCCACGATATCGATCTTTGGTGTTCGCTGTGTCATTCTAG	1320
Db	1261	GATAAGGTTGTTGCTATCCCCACGATATCGATCTTTGGTGTTCGCTGTGTCATTCTAG	1320
Qy	1321	GATTATGACAAACAGCATGAGCAGGATCGGCCACCCCGTGCCTGCGCCATCTCGGCT	1380
Db	1321	GATTATGACAAACAGCATGAGCAGGATCGGCCACCCCGTGCCTGCGCCATCTCGGCT	1380
Qy	1381	TTTTCTGTTCTCCGAGCAATGATGATCTTTGGTGTCTCCCTCACTGACCGAGTATGAC	1440
Db	1381	TTTTCTGTTCTCCGAGCAATGATGATCTTTGGTGTCTCCCTCACTGACCGAGTATGAC	1440
Qy	1441	CAGTCCACTTACGGTTCGTTTATATCTCGGACAGCGTGTGTTTGGT	1500
Db	1441	CAGTCCACTTACGGTTCGTTTATATCTCGGACAGCGTGTGTTTGGT	1500
Qy	1501	AATGTTGGAATGGCGCGAGCCGATAGCCGATCGTTGACTGGTCCAAAGTCAACCTC	1560
Db	1501	AATGTTGGAATGGCGCGAGCCGATAGCCGATCGTTGACTGGTCCAAAGTCAACCTC	1560
Qy	1561	GACGGCGGCCCTCCGACTGTTCAGCAATATTCAGAGATCTTTGTGTCTCCCTT	1620
Db	1561	GACGGCGGCCCTCCGACTGTTCAGCAATATTCAGAGATCTTTGTGTCTCCCTT	1620
Qy	1621	CGTGGCAAGCTCTCTTTTGGAGCGCGCAACAAAGAGAGGTTATCTTTATATAT	1680
Db	1621	CGTGGCAAGCTCTCTTTTGGAGCGCGCAACAAAGAGAGGTTATCTTTATATAT	1680
Qy	1681	AATACTAGTGTAGTGAACAGATCTGATTTGAATGCTGCGGCCATCGGTCGCATT	1740


```
Db 1501 AATGTTGCGACTGGCGCGCAGGCGGTAGCCGATCGCTTGACTGGTCCAAAGTCACCCCTC 1560
Qy 1561 GACGGCGGCCCTCCCGACTGTTGAGCAATATTCCAAAGCAATCTTTGTTGCTCCCCCTT 1620
Db 1561 GACGGCGGCCCTCCCGACTGTTGAGCAATATTCCAAAGCAATCTTTGTTGCTCCCCCTT 1620
Qy 1621 CGTGGCAAGCTCTCTCTTTTGGAGGCGCGCACAAACAAAGCAGGTTATCTTTAATTAAT 1680
Db 1621 CGTGGCAAGCTCTCTCTTTTGGAGGCGCGCACAAACAAAGCAGGTTATCTTTAATTAAT 1680
Qy 1681 AATACTACTGCTAGTGACAGATTTCTGATTGAAATGCTGCCGCCATCGGGTCGCCATT 1740
Db 1681 AATACTACTGCTAGTGACAGATTTCTGATTGAAATGCTGCCGCCATCGGGTCGCCATT 1740
Qy 1741 TCAACCTATACCAAGGCTTGGGCGGCTCGGTCGCCATTTCTGCGCGCGGGTTTG 1800
Db 1741 TCAACCTATACCAAGGCTTGGGCGGCTCGGTCGCCATTTCTGCGCGCGGGTTTG 1800
Qy 1801 GCTCCACGCTCCGCGCTGCTGCTGGAGGATACCTTTTGATTATCCGGGCGGGCGCAC 1860
Db 1801 GCTCCACGCTCCGCGCTGCTGCTGGAGGATACCTTTTGATTATCCGGGCGGGCGCAC 1860
Qy 1861 ACATTGATGACTTTCGCCCTGAATGCCCGCTTTAGGCTCCAGGGTTGCTTCCAG 1920
Db 1861 ACATTGATGACTTTCGCCCTGAATGCCCGCTTTAGGCTCCAGGGTTGCTTCCAG 1920
Qy 1921 TCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAGGTGAGTAAAGTCCGGAGTTGTAG 1980
Db 1921 TCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAGGTGAGTAAAGTCCGGAGTTGTAG 1980

RESULT 5
LOCUS AR002159 2100 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 2 from patent US 5741490.
ACCESSION AR002159
VERSION AR002159.1 GI:3963713
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Reyes,G.R., Bradley,D.W., Twu,J.-S., Purdy,M.A., Tam,A.W.,
Krawczynski,K.Z. and Yarbough,P.D.
TITLE Hepatitis B virus vaccine and method
JOURNAL Patent: US 5741490-A 2 21-APR-1998;
FEATURES
source Location/Qualifiers
1..2100
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 1980; DB 6; Length 2100;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGCCCTAGGCTCTTTGCTGTTGTTCTCTCTTTGCTGCTATGTTGCCGCGCA 60
Db 46 ATGCGCCCTAGGCTCTTTGCTGTTGTTCTCTCTTTGCTGCTATGTTGCCGCGCA 105
Qy 61 CCGACCGGTGACCGGTTGATTTCTAGCCCTTCCGAGTCCGAGCGGCGGTATCCGCGGTGT 120
Db 106 CCGACCGGTGACCGGTTGATTTCTAGCCCTTCCGAGTCCGAGCGGCGGTATCCGCGGTGT 165
Qy 121 TTCTGGGGTACCGGGTTGATTTCTAGCCCTTCCGAGTCCGAGTCCGAGTCCGAGTCC 180
Db 166 TTCTGGGGTACCGGGTTGATTTCTAGCCCTTCCGAGTCCGAGTCCGAGTCCGAGTCC 225
Qy 181 CCCTTTGCCCGCCAGACGTTGCGGCTCGGCTTCCGAGTCCGAGTCCGAGTCCGAGTCC 240
Db 226 CCCTTTGCCCGCCAGACGTTGCGGCTCGGCTTCCGAGTCCGAGTCCGAGTCCGAGTCC 285
Qy 241 CGGCCACTTGGGTTCCACTTGGCGAGATCAGGCCCGAGCGCCCTTCGCTGCCCTCCGCTCGC 300
```

```
Db 286 CGGCGCACTTGGCTCCACTTTGGCGAGATCAGGCCCGCCCTCGCTCGCTCCCGTCGC 345
Qy 301 CGACCTGCCACAGCGCGGCTCGCGGCTGAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCA 360
Db 346 CGACCTGCCACAGCGCGGCTCGCGGCTGAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCA 405
Qy 361 CCGGTCGCCGAGGTTGATTCTCGCGGTGCAATTTCTACGCCGCGCAGTATATTTGCTACT 420
Db 406 CCGGTCGCCGAGGTTGATTCTCGCGGTGCAATTTCTACGCCGCGCAGTATATTTGCTACT 465
Qy 421 TCACCCCTGACATCTCTCTGTGGCTCTTGCACTAATTTAGTCTCTGTATGAGCCCCCTT 480
Db 466 TCACCCCTGACATCTCTCTGTGGCTCTTGCACTAATTTAGTCTCTGTATGAGCCCCCTT 525
Qy 481 AATCGGCTCTCGCGCTGCGGCTGAGGAGGCTACTATACTACATATTAGGCCACAGAGGCTCC 540
Db 526 AATCGGCTCTCGCGCTGCGGCTGAGGAGGCTACTATACTACATATTAGGCCACAGAGGCTCC 585
Qy 541 AATTATGACAGTACCGGTTGCCCGCTACTATCCGTTACGGGCCCTAGTGCCTAAT 600
Db 586 AATTATGACAGTACCGGTTGCCCGCTACTATCCGTTACGGGCCCTAGTGCCTAAT 645
Qy 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTTTGGCCCTCAAAACAACCAACCCCTACA 660
Db 646 GCAGTTGGAGGCTATGCTATATCCATTTCTTTTGGCCCTCAAAACAACCAACCCCTACA 705
Qy 661 TCTGTTGACATGAATTCATTTACTTCCACTGATGTCAGGATTTCTGTTCAACCTGGCATA 720
Db 706 TCTGTTGACATGAATTTCCATTTCTTCCACTGATGTCAGGATTTCTGTTCAACCTGGCATA 765
Qy 721 GCATCTGAAATGGTCAATCCAAAGCGGCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
Db 766 GCATCTGAAATGGTCAATCCAAAGCGGCTTCACTACCGCAATCAAGGTTGGCGCTCG 825
Qy 781 GTTGAGACATCTGGTGTGCTGAGGAGGAGGCACTCCGGTCTTGTCTAATGTTATGCAFA 840
Db 826 GTTGAGACATCTGGTGTGCTGAGGAGGAGGCACTCCGGTCTTGTCTAATGTTATGCAFA 885
Qy 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTCGCCCTTGGCTTACTG 900
Db 886 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTCGCCCTTGGCTTACTG 945
Qy 901 GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAACACCAATACACCTGTG 960
Db 946 GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAACACCAATACACCTGTG 1005
Qy 961 TCCGTTACTCCAGCACTGCTGTCATCTCCGCCGAGGCGGCGACGCGGACTGCGGAGCTG 1020
Db 1006 TCCGTTACTCCAGCACTGCTGTCATCTCCGCCGAGGCGGCGACGCGGACTGCGGAGCTG 1065
Qy 1021 ACCCAACTGCGAGCCACCAAGTTTCATGAAGATCTCCACTTTTACCGGCCCTTAATGGGTA 1080
Db 1066 ACCCAACTGCGAGCCACCAAGTTTCATGAAGATCTCCACTTTTACCGGCCCTTAATGGGTA 1125
Qy 1081 GGTGAAGTCCGCGCGGATAGCTCTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1140
Db 1126 GGTGAAGTCCGCGCGGATAGCTCTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1185
Qy 1141 GGGCTCCCGACAGAAATTAATTTCTGCTGCGTGGGCGCAACTGTTTAACTTAACTTAACTTAACT 1200
Db 1186 GGGCTCCCGACAGAAATTAATTTCTGCTGCGTGGGCGCAACTGTTTAACTTAACTTAACTTAACT 1245
Qy 1201 GTCTCAGCAATGGCGAGCCACCGTGAAGCTCTATACATCAGTGGAGAACTCTCAGCAG 1260
Db 1246 GTCTCAGCAATGGCGAGCCACCGTGAAGCTCTTATCATCAGTGGAGAACTCTCAGCAG 1305
Qy 1261 GATAAGGTTGTTGCTATCCCCCAGATATCGATCTTTGTTGATTTCGGTGTGTTCAATCAG 1320
Db 1306 GATAAGGTTGTTGCTATCCCCCAGATATCGATCTTTGTTGATTTCGGTGTGTTCAATCAG 1365
Qy 1321 GATTATGACAAACAGCATGAGGAGATCGGCCCAACCCCTCGCTCGGCTCGGCTCGGCTCGGCT 1380
```


Db 1186 GGGCTCCGACAGAAATTAATTTTCGTGGCTGGCGGCAACTGTTTTATTTCGCCCGCGGTT 1245
QY 1201 GTCTCAGCAATGGCGAGCAACCGTGAAGCTCTATACATCAGTGGAGAAATCTCAGCAG 1260
Db 1246 GTCTCAGCAATGGCGAGCAACCGTGAAGCTCTATACATCAGTGGAGAAATCTCAGCAG 1305
QY 1261 GATAAGGGTGTGCTATCCCCACAGATATCGATCTTGGTGATTCGGCTGTGGTCAATTCAG 1320
Db 1306 GATAAGGGTGTGCTATCCCCACAGATATCGATCTTGGTGATTCGGCTGTGGTCAATTCAG 1365
QY 1321 GATTATGACAAACAGCAGCAGATCGGCCACACCCCGTCCGCTGCGGCATCTGGGCT 1380
Db 1366 GATTATGACAAACAGCAGCAGATCGGCCACACCCCGTCCGCTGCGGCATCTGGGCT 1425
QY 1381 TTTTCTGTTCTCGAGCAATATGATGATCTTTGGCTGTCCTCACTGCAGCCGAGTATGAC 1440
Db 1426 TTTTCTGTTCTCGAGCAATATGATGATCTTTGGCTGTCCTCACTGCAGCCGAGTATGAC 1485
QY 1441 CAGTCCACTTACGGGTGCGTCAACTGGCCCGGTTTATATCTCGGACAGCGGTGACTTTGGTG 1500
Db 1486 CAGTCCACTTACGGGTGCGTCAACTGGCCCGGTTTATATCTCGGACAGCGGTGACTTTGGTG 1545
QY 1501 AATGTTGCACTGGCGCGCAGCGCTGAGCCGATCGCTTGAATGCTGCCAAAGTCAACCTC 1560
Db 1546 AATGTTGCACTGGCGCGCAGCGCTGAGCCGATCGCTTGAATGCTGCCAAAGTCAACCTC 1605
QY 1561 GAGGGCGGCCCTCCGACTGTTGAGCAATATTCAGACATCTTTGCTGCTCCGCTT 1620
Db 1606 GAGGGCGGCCCTCCGACTGTTGAGCAATATTCAGACATCTTTGCTGCTCCGCTT 1665
QY 1621 CGTGGCAAGCTCTCCTTTTGGGAGCGCGGCACAAACAAAGCAGGTTATCTTATAATTAAT 1680
Db 1666 CGTGGCAAGCTCTCCTTTTGGGAGCGCGGCACAAACAAAGCAGGTTATCTTATAATTAAT 1725
QY 1681 AATACTACTGCTAGTGACAGATTTCTGATGAAATGCTGCCGCGCATCGGFTCGCCATT 1740
Db 1726 AATACTACTGCTAGTGACAGATTTCTGATGAAATGCTGCCGCGCATCGGFTCGCCATT 1785
QY 1741 TCAACCTATACCAACAGGCTTGGGCGGCTCGGTCGCCATTTCTGCGCGCGGFTTTG 1800
Db 1786 TCAACCTATACCAACAGGCTTGGGCGGCTCGGTCGCCATTTCTGCGCGCGGFTTTG 1845
QY 1801 GCTCCACGCTCGGCCCTCGCTGCTGAGAGTACTTTTGAATATCCGGGCGGCGGCAC 1860
Db 1846 GCTCCACGCTCGGCCCTCGCTGCTGAGGATCTTTTGAATATCCGGGCGGCGGCAC 1905
QY 1861 ACATTTGATGACTTCTGCTGCTGATCGCGCTTTAGGCTTCAGGGTTGCTTTCCAG 1920
Db 1906 ACATTTGATGACTTCTGCTGCTGATCGCGCTTTAGGCTTCAGGGTTGCTTTCCAG 1965
QY 1921 TCAACTGCTGCTGAGCTCCAGCGCCTTAAAGTTAAGTGGGTAAAACTCGGAGTTGTAG 1980
Db 1966 TCAACTGCTGCTGAGCTCCAGCGCCTTAAAGTTAAGTGGGTAAAACTCGGAGTTGTAG 2025

RESULT 7
LOCUS ARI150916 7171 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 10 from patent US 6229005.
ACCESSION ARI150916
VERSION ARI150916.1 GI:15115507
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 7171)
Reyes,G.R., Yarbough,P.O., Bradley,D.W., Krawczynski,K.Z., Tam,A.W.
and Fry,K.E.
TITLE DNA sequences of enterically transmitted non-A/non-B hepatitis
viral agent
JOURNAL Patent: US 6229005-A 10 08-MAY-2001;
FEATURES Location/Qualifiers
1..7171 source

/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 1980; DB 6; Length 7171;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTAGGCGCTTTTGGCTGTTGTTCTTCCTTGTGTTCTTCCTATGTTGCGCGGCA 60
Db 5117 ATGCGCCCTAGGCGCTTTTGGCTGTTGTTCTTCCTTGTGTTCTTCCTATGTTGCGCGGCA 5176
QY 61 CGACCGGTCAGCGCTCTGCGCGCGCTCTGCGCGCGCGCGCGCGGTACCGCGCGTGT 120
Db 5177 CGACCGGTCAGCGCTCTGCGCGCGCTCTGCGCGCGCGCGCGGTACCGCGCGTGT 5236
QY 121 TTCTGGGTCAGCGGTTGATTTCTCAGCGCTTCGCAATCCCTATATTCATCCAAACAAC 180
Db 5237 TTCTGGGTCAGCGGTTGATTTCTCAGCGCTTCGCAATCCCTATATTCATCCAAACAAC 5296
QY 181 CCTTTTGGCCCAAGACGTTTGGCTGCTGGCTCTGGACCTCGCCTTCGCCAACCGAGCC 240
Db 5297 CCTTTTGGCCCAAGACGTTTGGCTGCTGGCTCTGGACCTCGCCTTCGCCAACCGAGCC 5356
QY 241 CGGCCACTTTGGCTTCCACTTTGGCGAGATCAGGCCCGAGCGCCCTCGCTCGCTCGCTCGC 300
Db 5357 CGGCCACTTTGGCTTCCACTTTGGCGAGATCAGGCCCGAGCGCCCTCGCTCGCTCGCTCGC 5416
QY 301 CGACCTGCCACAGCGCGGCTCGCGGCTGAGCGCTGCGCGCTGCGCGCTGCGCCATGACACTCA 360
Db 5417 CGACCTGCCACAGCGCGGCTCGCGGCTGAGCGCTGCGCGCTGCGCGCTGCGCCATGACACTCA 5476
QY 361 CCGCTCCCGGAGCTTGATTTCTCGGTCGAATTTCTACGCGCGCAGTATATTTGTCTACT 420
Db 5477 CCGCTCCCGGAGCTTGATTTCTCGGTCGAATTTCTACGCGCGCAGTATATTTGTCTACT 5536
QY 421 TCACCCCTGACATCCTCTGTGGCTCTGGCACTAAATTTAGTCTGTATGAGCGCCCTT 480
Db 5537 TCACCCCTGACATCCTCTGTGGCTCTGGCACTAAATTTAGTCTGTATGAGCGCCCTT 5596
QY 481 AATCGCCTCTCGGCTCGAGACGCTACTATACATTCATATGCGCACAGAGGCTCC 540
Db 5597 AATCGCCTCTCGGCTCGAGACGCTACTATACATTCATATGCGCACAGAGGCTCC 5656
QY 541 AATTATGACAGTACCGGCTTGGCGCGCTACTATCCGTTACCGGCTTACGCGCCCTAGTGCCTAAT 600
Db 5657 AATTATGACAGTACCGGCTTGGCGCGCTACTATCCGTTACGCGCCCTAGTGCCTAAT 5716
QY 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTTCTGGCCTCAAAACAACACCCCTACA 660
Db 5717 GCAGTTGGAGGCTATGCTATATCCATTTCTTTCTGGCCTCAAAACAACACCCCTACA 5776
QY 661 TCTGTTGACATGAATTCATTTCTTCCACTGATGTCAGGATTTCTGTTCAACCTGGCATA 720
Db 5777 TCTGTTGACATGAATTCATTTCTTCCACTGATGTCAGGATTTCTGTTCAACCTGGCATA 5836
QY 721 GCATCTGAATTTGGTCAATCCCAAGCGAGCGCTTCACTACCGCAATCAAGGTTTGGCGCTCG 780
Db 5837 GCATCTGAATTTGGTCAATCCCAAGCGAGCGCTTCACTACCGCAATCAAGGTTTGGCGCTCG 5896
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGGAGCGCATCTCGGCTTCATGTTATGCTATG 840
Db 5897 GTTGAGACATCTGGTGTGCTGAGGAGGAGGAGCGCATCTCGGCTTCATGTTATGCTATG 5956
QY 841 CATGGCTCTCCAGTTTAACTTCACTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 900
Db 5957 CATGGCTCTCCAGTTTAACTTCACTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 6016
QY 901 GACTTTGCTTGAAGCTTGAGTTTTCGCAATCTCACACCTGTAAACAACCAATACACGTTG 960
Db 6017 GACTTTGCTTGAAGCTTGAGTTTTCGCAATCTCACACCTGTAAACAACCAATACACGTTG 6076
QY 961 TCCCGTTACTCCAGCACTGCTGCTACTCGCGCCGAGGGGCCACCGGGAGCTGCGGAGCTG 1020

Db 5357 CGGCCACTTTGGCTCCACTTTGGCGAGATCAGGCCAGCGCCCTCCGCTCCCTCCGTCGC 5416
QY 301 CGACCTGCCACAGCGGGGCTCGCGGCTGACGGCTGTGGCGCCTGCCCATGACACCTCA 360
Db 5417 CGACCTGCCACAGCGGGGCTCGCGGCTGACGGCTGTGGCGCCTGCCCATGACACCTCA 5476
QY 361 CCCGTCGCCGAGCGTTGATTCTCGCGGTGCAATTTCTACGCCGCGAGTATAAATTTGCTACT 420
Db 5477 CCCGTCGCCGAGCGTTGATTCTCGCGGTGCAATTTCTACGCCGCGAGTATAAATTTGCTACT 5536
QY 421 TCACCCCTGACATCCTCTGTGGCTCTGSCACTAATTTAGTCCTGTAAGCAGCCGCCCTT 480
Db 5537 TCACCCCTGACATCCTCTGTGGCTCTGSCACTAATTTAGTCCTGTAAGCAGCCGCCCTT 5596
QY 481 AATCCGCTCTCCGCTGCGAGACGCTACTAATCTCACATTAATGSCACAGAGCCTCC 540
Db 5597 AATCCGCTCTCCGCTGCGAGACGCTACTAATCTCACATTAATGSCACAGAGCCTCC 5656
QY 541 AATTATGACAGTACCGGTTGCGCGCTACTATCCGTTACCGGCCCTTAGTGCTTAAT 600
Db 5657 AATTATGACAGTACCGGTTGCGCGCTACTATCCGTTACCGGCCCTTAGTGCTTAAT 5716
QY 601 GCAGTTGAGGCTATGCTATATCAATTTCTTGGCTCAAAACACCAACCCCTACA 660
Db 5717 GCAGTTGAGGCTATGCTATATCAATTTCTTGGCTCAAAACACCAACCCCTACA 5776
QY 661 TCTGTTGACATGAATTCATTTCACTGATGTCAGGATCTTTGTTCAACCTGGCATA 720
Db 5777 TCTGTTGACATGAATTCATTTCACTGATGTCAGGATCTTTGTTCAACCTGGCATA 5836
QY 721 GCATCTGAATTTGATCATCCCAAGCGAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
Db 5837 GCATCTGAATTTGATCATCCCAAGCGAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 5896
QY 781 GTTGAGATCTGGGTGCTGAGGAGGAGCCACTCGGCTTGTGATGATGATGATA 840
Db 5897 GTTGAGATCTGGGTGCTGAGGAGGAGCCACTCGGCTTGTGATGATGATGATA 5956
QY 841 CATGCTCTCCAGTTAACTCCCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 900
Db 5957 CATGCTCTCCAGTTAACTCCCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 6016
QY 901 GACTTTGCTTAGAGCTTTGAGTTTCGCAATCTCACCACTGTAAACCAATACAGTGTG 960
Db 6017 GACTTTGCTTAGAGCTTTGAGTTTCGCAATCTCACCACTGTAAACCAATACAGTGTG 6076
QY 961 TCCGTTTACTCAGACCTGCTGTCACCTCGGCCGAGGGGCGGAGCTGCGAGCTG 1020
Db 6077 TCCGTTTACTCAGACCTGCTGTCACCTCGGCCGAGGGGCGGAGCTGCGAGCTG 6136
QY 1021 ACCACAACTGCAGCCACCAAGTTCATGAAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080
Db 6137 ACCACAACTGCAGCCACCAAGTTCATGAAAGATCTCCACTTTACCGGCTTAAATGGGTA 6196
QY 1081 GGTGAAGTTCGCGCGGATAGCTCTAATTAATTAACCTTGTGACACGCTCTCGGCG 1140
Db 6197 GGTGAAGTTCGCGCGGATAGCTCTAATTAATTAACCTTGTGACACGCTCTCGGCG 6256
QY 1141 GGGCTCCGACAGATTAATTTGCTGCGCTGCGGCGAATGTTTTATCCCGCGGCTT 1200
Db 6257 GGGCTCCGACAGATTAATTTGCTGCGCTGCGGCGAATGTTTTATCCCGCGGCTT 6316
QY 1201 GTCTCAGCAATGGCGAGCCCAACCGTGAAGTCTATACATCAGTGGAGAATGCTCAGAG 1260
Db 6317 GTCTCAGCAATGGCGAGCCCAACCGTGAAGTCTATACATCAGTGGAGAATGCTCAGAG 6376
QY 1261 GATAAGGCTGTGCTATCCCCCAGATATCGATCTTGGTATTCGCGTGTGGTCAATTCAG 1320
Db 6377 GATAAGGCTGTGCTATCCCCCAGATATCGATCTTGGTATTCGCGTGTGGTCAATTCAG 6436
QY 1321 GATTATGACACACGATGACAGGATCGGCGCACCCCGCTCGCTGCGCCATCTCGGCT 1380
Db 6437 GATTATGACACACGATGACAGGATCGGCGCACCCCGCTCGCTGCGCCATCTCGGCT 6496

QY 1381 TTTTCTGTTCTCGGAGCAATGATGATCTTTGGCTGTCCCTCACTGCACGCGATGATGAC 1440
Db 6497 TTTTCTGTTCTCGGAGCAATGATGATCTTTGGCTGTCCCTCACTGCACGCGATGAC 6556
QY 1441 CAGTCCACTTACGGGTGCTCAACTGCGCCGGTTTATATCTCGGACAGCGTGAATTTGGTG 1500
Db 6557 CAGTCCACTTACGGGTGCTCAACTGCGCCGGTTTATATCTCGGACAGCGTGAATTTGGTG 6616
QY 1501 AATGTTGCACTGCGCGGAGGCCGATGAGCCGATCGCTTGACTGTGTCACCAAGTCAACCTTC 1560
Db 6617 AATGTTGCACTGCGCGGAGGCCGATGAGCCGATCGCTTGACTGTGTCACCAAGTCAACCTTC 6676
QY 1561 GACGGGCGGCCCTCCGACTGTTGAGCAATATTTCAAGACATTTCTTGTGTCGCCCTT 1620
Db 6677 GACGGGCGGCCCTCCGACTGTTGAGCAATATTTCAAGACATTTCTTGTGTCGCCCTT 6736
QY 1621 CTTGCAAGCTCTCTCTTTTGGAGCGCGGACACAAAGAGAGGTTATCTTATTAATAT 1680
Db 6737 CTTGCAAGCTCTCTCTTTTGGAGCGCGGACACAAAGAGAGGTTATCTTATTAATAT 6796
QY 1681 AATACTACTGCTAGTGACAGATTTCTGATGAAATGCTGCCGCCATCGGCTGCCATTT 1740
Db 6797 AATACTACTGCTAGTGACAGATTTCTGATGAAATGCTGCCGCCATCGGCTGCCATTT 6856
QY 1741 TCAACCTATACACAGCGTTGGGCGCGTCCGCTCGGCAATTTCTCGGCCCGCGTTTG 1800
Db 6857 TCAACCTATACACAGCGTTGGGCGCGTCCGCTCGGCAATTTCTCGGCCCGCGTTTG 6916
QY 1801 GTCACACCTCGGCCCTCGCTCTGCTGGAGGATATCTTTGATTATCCGGGCGGCGGCAC 1860
Db 6917 GTCACACCTCGGCCCTCGCTCTGCTGGAGGATATCTTTGATTATCCGGGCGGCGGCAC 6976
QY 1861 ACATTTGATGATCTTCTGCCCTGGAATGCGCGCTTTAGGCTCCAGGTTGTCTTCCAG 1920
Db 6977 ACATTTGATGATCTTCTGCCCTGGAATGCGCGCTTTAGGCTCCAGGTTGTCTTCCAG 7036
QY 1921 TCAACTGTCGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTGTAG 1980
Db 7037 TCAACTGTCGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTGTAG 7096

RESULT 10

HPENSSP
LOCUS Hepatitis E virus (mexican strain) linear VRL 02-AUG-1993
DEFINITION nonstructural polyprotein genes, complete cds.
ACCESSION M74506
VERSION 1
KEYWORDS nonstructural protein; structural protein.
SOURCE Hepatitis E virus
ORGANISM Hepatitis E virus
REFERENCE 1. (bases 1 to 7180)
AUTHORS Huang C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E.,
Bradley D.W., Tam A.W. and Reyes G.R.
TITLE Molecular cloning and sequencing of the Mexico isolate of hepatitis
E virus (HEV)
JOURNAL Virology 191 (2), 550-558 (1992)
MEDLINE 93079857
PUBMED 1448913
COMMENT Original source text: Hepatitis E virus cDNA to genomic RNA.
FEATURES
source
1. 7180
/organism="Hepatitis E virus"
/mol_type="genomic RNA"
/db_xref="taxon:12461"
4. 5079
/note="putative"
/codon_start=1
/product="nonstructural polyprotein"
/protein_id="AA045730.1"
CDS

Db 6317 GTCTCAGCAATGGCGAGCCACCGTGAAGCTCTATACATCATGTGAGAAATGCTCAGCAG 6376
QY 1261 GATAAGGGGTGTGCTATCCGCCACGATATCGATCTTGTGATTCGGTGTGTCATTACAG 1320
Db 6377 GATAAGGGGTGTGCTATCCGCCACGATATCGATCTTGTGATTCGGTGTGTCATTACAG 6436
QY 1321 GATTATGACAAACAGATGAGCAGGATCGGCCACCCCGCTGCGCCATCTCGGCCCT 1380
Db 6437 GATTATGACAAACAGATGAGCAGGATCGGCCACCCCGCTGCGCCATCTCGGCCCT 6496
QY 1381 TTTTCTGTCTCCGAGCAATGATGATCTTTGGCTGTGCTCAGTCCGAGCTGATGAC 1440
Db 6497 TTTTCTGTCTCCGAGCAATGATGATCTTTGGCTGTGCTCAGTCCGAGCTGATGAC 6556
QY 1441 CAGTCCACTTACGGGTGCTCAACTGGCCGGTTTATATCTCGGACAGCTGACTTTGGTG 1500
Db 6557 CAGTCCACTTACGGGTGCTCAACTGGCCGGTTTATATCTCGGACAGCTGACTTTGGTG 6616
QY 1501 AATGTTGCACTGGCGGAGCCGCTAGCCCGATCGCTTGACTGGTCCAAAGTCAACCCTC 1560
Db 6617 AATGTTGCACTGGCGGAGCCGCTAGCCCGATCGCTTGACTGGTCCAAAGTCAACCCTC 6676
QY 1561 GACGGCGGCCCTCCGACCTGTTGAGCAATATTCAGACATCTTTGTGTCGCCCTT 1620
Db 6677 GACGGCGGCCCTCCGACCTGTTGAGCAATATTCAGACATCTTTGTGTCGCCCTT 6736
QY 1621 CGTGGCAAGCTCTCCTTTTGGAGCGCGGCAACAAAGACAGGTTATCCCTATATATAT 1680
Db 6737 CGTGGCAAGCTCTCCTTTTGGAGCGCGGCAACAAAGACAGGTTATCCCTATATATAT 6796
QY 1681 AATACTACTGCTAGTACAGATCTGATTTGAAATGCTCGCGGCATCGGCTCGCCATT 1740
Db 6797 AATACTACTGCTAGTACAGATCTGATTTGAAATGCTCGCGGCATCGGCTCGCCATT 6856
QY 1741 TCAACCTATACACAGGCTTGGGCGCGCTCGGTCGCAATTTTGGCGCGCGGTTTG 1800
Db 6857 TCAACCTATACACAGGCTTGGGCGCGCTCGGTCGCAATTTTGGCGCGCGGTTTG 6916
QY 1801 GCTCACGCTCGGCCCTCGGCTCTGCTGGAGATACCTTTTGATATATCCGGGCGGCGCAC 1860
Db 6917 GCTCACGCTCGGCCCTCGGCTCTGCTGGAGATACCTTTTGATATATCCGGGCGGCGCAC 6976
QY 1861 ACATTTGATGACTTGTGCTTGAATCGCGCTTTAGGCTTCCAGGGTTGTGCTTCCAG 1920
Db 6977 ACATTTGATGACTTGTGCTTGAATCGCGCTTTAGGCTTCCAGGGTTGTGCTTCCAG 7036
QY 1921 TCAACTGCTGCTGAGCTCCAGGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTGTAG 1980
Db 7037 TCAACTGCTGCTGAGCTCCAGGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTGTAG 7096

RESULT 11
AR170407
LOCUS AR170407 1644 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 4 from patent US 6291641.
ACCESSION AR170407
VERSION AR170407.1 GI:17908366
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1644)
AUTHORS Fuerst,T.R., McAtee,C.Patrick., Yarbough,P.O. and Zhang,Y.-F.
TITLE Hepatitis E virus antigens and uses therefor
JOURNAL Patent: US 6291641-A 4 18-SEP-2001;
FEATURES
Location/Qualifiers
source 1..1644
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 83.0%; Score 1644; DB 6; Length 1644;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1644; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GCTGTGGGCGCTGCGCATGACACCTCAACCGTCCCGGACGTTGATTTCTCGCGGTGCAATT 393
Db 1 GCTGTGGGCGCTGCGCATGACACCTCAACCGTCCCGGACGTTGATTTCTCGCGGTGCAATT 60
QY 394 CTACGCGGCGAGTATAAATTTGTCTACTTCCACCCCTGACATCCTCTGTGGCCTCTGCGACT 453
Db 61 CTACGCGGCGAGTATAAATTTGTCTACTTCCACCCCTGACATCCTCTGTGGCCTCTGCGACT 120
QY 454 AATTTAGTCTGTATGACAGCCCCCTTAATCGCCCTCTGCGCCTGCGAGACGGTACTAAT 513
Db 121 AATTTAGTCTGTATGACAGCCCCCTTAATCGCCCTCTGCGCCTGCGAGACGGTACTAAT 180
QY 514 ACTCAGATTTGGCCACAGAGGCTCCAAATATGACAGTACCGGTTGCGCGCTACT 573
Db 181 ACTCAGATTTGGCCACAGAGGCTCCAAATATGACAGTACCGGTTGCGCGCTACT 240
QY 574 ATCCGTTTACCGGCCCTTAGTGCTTAATGCAGTTGAGGCTATGCTATATCCATTTCTTTC 633
Db 241 ATCCGTTTACCGGCCCTTAGTGCTTAATGCAGTTGAGGCTATGCTATATCCATTTCTTTC 300
QY 634 TGGCTCTAAACACACAAACCCCTACATCTGTGACATGAATTCATTTACTTCCACTGAT 693
Db 301 TGGCTCTAAACACACAAACCCCTACATCTGTGACATGAATTCATTTACTTCCACTGAT 360
QY 694 GTCAGGATTTCTTTCAACCTGGGATAGCATCTGAATTTGGTCAATCCAGAGGCGCTT 753
Db 361 GTCAGGATTTCTTTCAACCTGGGATAGCATCTGAATTTGGTCAATCCAGAGGCGCTT 420
QY 754 CACTACCGCAATCAAGTTGGGCTCGGTTGAGCATCTGGTTGGTCTGAGGAGGAAAGCC 813
Db 421 CACTACCGCAATCAAGTTGGGCTCGGTTGAGCATCTGGTTGGTCTGAGGAGGAAAGCC 480
QY 814 ACCTCCGCTCTTGTATGATATGATATGATGCTCCAGTTAACTCTATACCAATACC 873
Db 481 ACCTCCGCTCTTGTATGATATGATATGATGCTCCAGTTAACTCTATACCAATACC 540
QY 874 CTTTATACCGTGGCTTTGGCTTACTGGAATTTGCTTAGAGCTTGAATTTGCAATCTC 933
Db 541 CTTTATACCGTGGCTTTGGCTTACTGGAATTTGCTTAGAGCTTGAATTTGCAATCTC 600
QY 934 ACCACTGTAAACCAATACACGCTGTCCGCTTACTCCAGCATGCTCGTCACTCCGCC 993
Db 601 ACCACTGTAAACCAATACACGCTGTCCGCTTACTCCAGCATGCTCGTCACTCCGCC 660
QY 994 CGAGGGCGGACGGGACTGCGGAGCTGACCAAACTGCAGCCACCGAGTTCAATGAAGAT 1053
Db 661 CGAGGGCGGACGGGACTGCGGAGCTGACCAAACTGCAGCCACCGAGTTCAATGAAGAT 720
QY 1054 CTCACATTTACCGGCTTAAATGAGGTAGTGAAGTGGCGCGGGATAGCTCTTAACATTA 1113
Db 721 CTCACATTTACCGGCTTAAATGAGGTAGTGAAGTGGCGCGGGATAGCTCTTAACATTA 780
QY 1114 CTTAAACCTTGTGACACGCTCTCGGCGGCTCCCGGACAGAAATTAATTTCTCGCTGGC 1173
Db 781 CTTAAACCTTGTGACACGCTCTCGGCGGCTCCCGGACAGAAATTAATTTCTCGCTGGC 840
QY 1174 GGGCAACTGTTTTATTTCCCGCCCGGTTGTCTCAGCCAAATGGCGAGCCAAACCGTGAAGCT 1233
Db 841 GGGCAACTGTTTTATTTCCCGCCCGGTTGTCTCAGCCAAATGGCGAGCCAAACCGTGAAGCT 900
QY 1234 TATACATCAGTGAGAAATGCTCAGCAGGATAAGGTTGTTGTAATCCCGGACGATATCGAT 1293
Db 901 TATACATCAGTGAGAAATGCTCAGCAGGATAAGGTTGTTGTAATCCCGGACGATATCGAT 960
QY 1294 CTTTGTGATTTGGGCTGTTGTTGATTCAGGATATGACAAACAGCATGAGCAGGATCGGCC 1353
Db 961 CTTTGTGATTTGGGCTGTTGTTGATTCAGGATATGACAAACAGCATGAGCAGGATCGGCC 1020
QY 1354 ACCCGGTCGCTGCGCCATCTCGGCCCTTTTCTGTTCTCCGAGCAATGATGACTTTGG 1413
Db 1021 ACCCGGTCGCTGCGCCATCTCGGCCCTTTTCTGTTCTCCGAGCAATGATGACTTTGG 1080


```

/codon_start=1
/protein_id="AAA45727.1"
/db_xref="GI:330000"
/translation="MRPEPILLLILFELPMLPAPPQPSGRRRRSGSGSGGFWGD
RVDSQFPALPIYHPNTPEFAPDVTAAAGAGPVRDPARLPDSAGRWDRQAPRAASRRRP
TTAAGAPLTAVAPAHDTPEVDVDSRGDAIRKQYNLSTSPLTSSVATGNTLVLYAAPL
SPLPLDGTNGHTNIMATASNYAQVRAATIRYPLPVNAPGVGAYISLSPMPQTIT
PTVDMNSITSDVRLVQPGFASLEVLPSERLHYRNCQWRVSVSGVAEEASTGLV
MLC.IHGSFVNSVYTPYKGLGALLDPALEFRNLTPGNTNTRVSRSTARRLRKG
GQDAEALLTTAATPRMKDLYFTSTNGVGEIRGIALTLFNALDTLGGITPELLISAG
GOLFVSRVNSANGBETVKLYISVENAQDKQIALPHDIDLGSRVYIQDYNQHQSD
RFTSPSPSPSVLRVLRANDVWLSTAAEYDOSTYGSSTPVGVSDSVTLFVNQTGAS
AVASRLDKVTFLDGRPLSTIQYQSKTFFVLPRGLKLSFWEAQTTKAGYVYNTTQAS
DQLLVENAGHRVATLSTYSLGAGPVSISAVAVLALLESVLLLEDTMDYPARAHTFD
DECPRCRPLGIGCGCAFOSTVAELORI.KMKVQSGTREI."

```

ORIGIN

[illegible]

Db	6527	CCCTTTTCTGTCCTCCGAGCTAACGATGTGCTTTGGCTTTCTCTCACCGTCCGAGTAT	6586
Qy	1438	GACCAGTCCACTTACGGGTGCTCAACTGGCCCGGTTTATATCTCGGACAGCTGACTTTG	1497
Db	6587	GACCAGTCCACTTACGGCTCTTCGACGGCCAGTCTATGCTCTGACTCTGTGACCTTG	6646
Qy	1498	GTGAATGTTGGGACTGGCGCGCAGCCGTAGCCGATCGCTTGACTGGTCCAAAGTCACC	1557
Db	6647	GTTTAATGTTGGACCGCGCGCAGCCGTTGCCCGGTCACTCGACTGGACCAAGGTACA	6706
Qy	1558	CTGACGGGGGGCCCTCCGACCTGTTGAGCAATATTCAAAGACATTCTTTGTGCTCCCC	1617
Db	6707	CTTGATGGTCGCCCCCTTCCACCATCCAGCAGTATTCAAAGACCTTCTTTGCTGCGG	6766
Qy	1618	CTTCGTGGCAAGCTCTCGTTTTGGGAGCCGGGCAACAACAAAGCAGGTTATCCTTATAAT	1677
Db	6767	CTCCGCGGTAAAGCTCTCCTTTTGGGAGGCGAGAACTACTAAGCCGGGTACCTCTATAAT	6826
Qy	1678	TATAATACTACTGCTAGTGACCAGATTCTGATTGAAAATGCTGCGGCCCATCGGGTCGCC	1737
Db	6827	TATAACACCACCTGCTAGTGACCAACTGCTCGTTGAGAAATCGCGCTGGGCATCGGTTGCT	6886
Qy	1738	ATTTCAACCTATACCAACAGCTTGGGGCCGGTCGGCTCGGCATTTCTGCGGCCGCGGTT	1797
Db	6887	ATTTTCCACCTACACTACTAGCTGGGTGCTGGCCCGGCTCTATATTTCCGGGTGCTGTT	6946
Qy	1798	TTGGCTCCAGCTCCGGCCCTGGCTCTGCTGGAGGATACTTTTGATTATCCGGGGCGGGG	1857
Db	6947	TTAGCCCCACACTCTGTGCTAGCATTGCTTGAGATACCATGGACTACCTGCCCCGCC	7006
Qy	1858	CACACATTTGATGACTTCTGCGCTTGAATCCGCGCTTTAGCCCTCCAGGGTGTGCTTTC	1917
Db	7007	CATATCTTCATGACTTCTGCGCCGAGTGGCGCCCGCTTGCCCTCCAGGGTGTGCTTTT	7066
Qy	1918	CAGTCAACTGTCGCTGAGCTCCAGCGCTTAAAGCTTAAGGTGGTAAACCTCGGAGCTG	1977
Db	7067	CAGTCTACTGTCGCTGAGCTTCAGCGCTTAAAGTGAAGGTGGTAAACCTCGGAGTTA	7126
Qy	1978	TAG 1980	
Db	7127	TAG 7129	

Search completed: August 21, 2004, 15:15:06
Job time : 7637.72 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 08:40:47 ; Search time 764.421 Seconds
(without alignments)
11003.666 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096
Perfect score: 1980
Sequence: 1 ATGGCGCCTAGGCCTCTTTT.....GTAAACTCGGAGTGTGAG 1980

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1980	100.0	1980	2	AAT96960	Aat96960 Hepatitis
2	1980	100.0	2100	2	AQ47130	Aq47130 HEV third
3	1980	100.0	7171	2	AAQ14413	Aaql4413 Composite
4	1980	100.0	7171	2	AAV66322	Aav66322 ET-NANB (
5	1980	100.0	7171	2	AAV54730	Aav54730 Composite
6	1980	100.0	7171	3	AA99260	Aa99260 Hepatitis
7	1980	100.0	7171	4	AAF83496	Aaf83496 Mexican s
8	1980	100.0	7171	6	AAAL50387	Aal50387 Hepatitis
9	1980	100.0	7171	9	ADD24378	Add24378 Hepatitis
10	1980	100.0	7180	9	ADE06698	Ade06698 Hepatitis
11	1978.4	99.9	2100	2	AQ46814	Aq46814 Mexico st
12	1957.4	98.9	2058	2	AAT27108	Aat27108 Hepatitis
13	1621.4	81.9	1647	2	AAT27110	Aat27110 Hepatitis
14	1378.2	69.6	1983	2	AAV61688	Aav61688 Hepatitis
15	1378.2	69.6	7158	2	AAT27394	Aat27394 Hepatitis
16	1378.2	69.6	7168	2	AQ45197	Aq45197 HEV strai
17	1378.2	69.6	7168	2	AAV71604	Aav71604 Hepatitis
18	1378.2	69.6	7194	2	AAV61690	Aav61690 Hepatitis
19	1378.2	69.6	7204	8	ADA50064	Ada50064 Hepatitis
20	1378.2	69.6	7204	8	ADA50062	Ada50062 SK-HEV-2
21	1376.6	69.5	7204	8	ADA50063	Ada50063 SK-HEV-3
22	1376.6	69.5	7204	8	ADA50065	Ada50065 Hepatitis
23	1371.8	69.3	1984	2	AAQ86592	Aaq86592 Hepatitis

24	1371.8	69.3	1990	6	ABL60053	Abi60053 Hepatitis
25	1371.8	69.3	1990	9	ADB97808	Adb97808 HEV ORF2
26	1371.8	69.3	2054	4	AA501448	Aa501448 Hepatitis
27	1368.6	69.1	1983	2	AAT96959	Aat96959 Hepatitis
28	1368.6	69.1	2049	2	AAT27107	Aat27107 Hepatitis
29	1368.6	69.1	2094	2	AAQ47129	Aaq47129 HEV third
30	1368.6	69.1	2094	2	AAQ46813	Aa46813 Burma str
31	1368.6	69.1	7195	2	AAV66321	Aav66321 ET-NANB (
32	1368.6	69.1	7195	2	AAV54729	Aav54729 DNA seque
33	1368.6	69.1	7195	3	AAV9259	Aav9259 Hepatitis
34	1368.6	69.1	7195	4	AAF83495	Aaf83495 ET-NANB v
35	1368.6	69.1	7195	6	AAAL50386	Aal50386 Hepatitis
36	1363.8	68.9	7195	9	ADD24374	Add24374 Hepatitis
37	1361.4	68.8	1989	6	ABL60054	Abi60054 Hepatitis
38	1361.4	68.8	1989	9	ADB97809	Adb97809 HEV relat
39	1355.8	68.5	7195	2	AAQ14412	Aaql4412 Forward s
40	1319.6	66.6	7256	9	ADE06693	Ade06693 Hepatitis
41	1315.2	66.4	7277	2	AAZ00267	Aaz00267 HEV-US2 O
42	1315.2	66.4	7277	2	AAZ00266	Aaz00266 HEV-US2 f
43	1314	66.4	7228	9	ADE06696	Ade06696 Hepatitis
44	1313.2	66.3	7233	9	ADE06694	Ade06694 Hepatitis
45	1301.2	65.7	7229	9	ADE06695	Ade06695 Hepatitis

ALIGNMENTS

RESULT 1
AAT96960
ID AAT96960 standard; DNA; 1980 BP.
XX
AC AAT96960;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 26-FEB-1998 (first entry)
XX
DE Hepatitis E virus Mexico strain ORF2.
XX
KW Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody;
KW diagnosis; HEV; ds.
XX
OS Hepatitis E virus; - Mexico strain.
FH Key Location/Qualifiers
FT CDS 1..1980
FT /*tag= a
FT /product= "ORF2_protein"
XX
PN US5686239-A.
XX
PD 11-NOV-1997.
XX
PF 09-MAY-1994; 9AUS-00240049.
XX
PR 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 85US-00336672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 05-APR-1991; 91US-00681078.
PR 17-JAN-1992; 92US-00822335.
PR 20-APR-1992; 92US-00870985.
PR 01-MAY-1992; 92US-00876941.
XX
(GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Tam AW, Reyes GR, Yarbrough PO;
XX
DR WPI; 1997-558132/51.
XX
P-PSDB; AAW35827.
PT Diagnosis of hepatitis E virus Burma and Mexico strain infection - by

PT immunoassay with hepatitis E virus peptide antigens.

PS Disclosure; Col 25-28, 36pp; English.

XX A method has been developed for detecting hepatitis E virus (HEV) antibodies (Ab). The method comprises: (a) reacting a serum sample with a HEV peptide antigen; and (b) examining the peptide for the presence of bound Ab, where the presence of bound Ab indicates the presence of HEV Ab. The present sequence encodes the protein from the open reading frame, ORF2, from HEV Mexico strain. The method can be used to diagnose infection with the enterically transmitted non-A/non-B viral hepatitis agent HEV, specifically the HEV Burma and Mexico strains. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 1980 BP; 364 A; 624 C; 493 G; 499 T; 0 U; 0 Other;

Query Match 100.0%; Score 1980; DB 2; Length 1980;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTAGGCTCTTTGCTGTGTTCCCTTGTCTGCTATGTTGCCGCGCA 60

DB 1 ATGGCCCTAGGCTCTTTGCTGTGTTCCCTTGTCTGCTATGTTGCCGCGCA 60

QY 61 CCGACCGGTACGCGCTCTGGCCGCGCTGTGGGCGCGCAGCGGTACCGCGGTGT 120

DB 61 CCGACCGGTACGCGCTCTGGCCGCGCTGTGGGCGCGCAGCGGTACCGCGGTGT 120

QY 121 TTCTGGGGTACCGGGTGTGATTTCTAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180

DB 121 TTCTGGGGTACCGGGTGTGATTTCTAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180

QY 181 CCCTTTGCCCGACAGCTTCGCGCTGCTGGGTCTGACCTCGCCTTCCCAACAGCC 240

DB 181 CCCTTTGCCCGACAGCTTCGCGCTGCTGGGTCTGACCTCGCCTTCCCAACAGCC 240

QY 241 CGGCCACTTGGCTTCCACTTGGGAGATCAGGCCAGCGCCCTCGCTCCGCTCGC 300

DB 241 CGGCCACTTGGCTTCCACTTGGGAGATCAGGCCAGCGCCCTCGCTCCGCTCGC 300

QY 301 CGACCTGCCACAGCGGGGCTGCGGCTGACGGCTGTGGGCTGTGCCATGACACTCA 360

DB 301 CGACCTGCCACAGCGGGGCTGCGGCTGACGGCTGTGGGCTGTGCCATGACACTCA 360

QY 361 CCGCTCCCGGAGCTTGAATCTCGCGTGCATTTCTACGCGCGCAGTATATTTGCTACT 420

DB 361 CCGCTCCCGGAGCTTGAATCTCGCGTGCATTTCTACGCGCGCAGTATATTTGCTACT 420

QY 421 TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTTAGTCTGTATGACGCCCTCTT 480

DB 421 TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTTAGTCTGTATGACGCCCTCTT 480

QY 481 AATCGGCTCTGCGCTGACAGCGGTACTAATCTACATATGCGCAAGAGGCTCC 540

DB 481 AATCGGCTCTGCGCTGACAGCGGTACTAATCTACATATGCGCAAGAGGCTCC 540

QY 541 AATTATGACAGTACCGGGTGTCCGCGTACTATCCGTTACCGGCCCTAGTGCCTAAT 600

DB 541 AATTATGACAGTACCGGGTGTCCGCGTACTATCCGTTACCGGCCCTAGTGCCTAAT 600

QY 601 CGAGTTGGAGGCTATGCTATATCCATTTCTTGTGGCTCAAAACCAACCCCTACA 660

DB 601 CGAGTTGGAGGCTATGCTATATCCATTTCTTGTGGCTCAAAACCAACCCCTACA 660

QY 661 TCTGTTGACATGAATTCATTTCTTCTGCTGATGTCAGGATTTCTGTTCAACTGSCATA 720

DB 661 TCTGTTGACATGAATTCATTTCTTCTGCTGATGTCAGGATTTCTGTTCAACTGSCATA 720

QY 721 GCATCTGAATTTGGTCAATCCCAAGCAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 780

DB 721 GCATCTGAATTTGGTCAATCCCAAGCAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 780

QY 781 GTTGAGACATCTGGTGTCTGCTGAGGAGAACACCTCGGCTCTTGTCTGATGTTATGCATA 840

DB 781 GTTGAGACATCTGGTGTCTGCTGAGGAGAACACCTCGGCTCTTGTCTGATGTTATGCATA 840

QY 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 900

DB 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 900

QY 901 GACTTTGCTTTAGAGCTTGAGTTTCCATCTCAACCTGTAAACACCAATACACCTGTG 960

DB 901 GACTTTGCTTTAGAGCTTGAGTTTCCATCTCAACCTGTAAACACCAATACACCTGTG 960

QY 961 TCCCGTTACTCCAGCACTGCTCGTCACTCCGCCGAGGGCGGACGGGAGCTGCGGAGCTG 1020

DB 961 TCCCGTTACTCCAGCACTGCTCGTCACTCCGCCGAGGGCGGACGGGAGCTGCGGAGCTG 1020

QY 1021 ACCAACAATGAGCCACAGGTTTCATGAAAGATCTCTCACTTTACCGGCTTAAATGGGGTA 1080

DB 1021 ACCAACAATGAGCCACAGGTTTCATGAAAGATCTCTCACTTTACCGGCTTAAATGGGGTA 1080

QY 1081 GGTGAAGTCGCGCGGGGATAGCTCTAACATTACTTAACTTGTCTGACACGCTCTCCGGC 1140

DB 1081 GGTGAAGTCGCGCGGGGATAGCTCTAACATTACTTAACTTGTCTGACACGCTCTCCGGC 1140

QY 1141 GGGCTCCCGACAGAAATTAATTTCTGCGCTGGCGGCAACTGTTTTATTTCCGCGCCGTT 1200

DB 1141 GGGCTCCCGACAGAAATTAATTTCTGCGCTGGCGGCAACTGTTTTATTTCCGCGCCGTT 1200

QY 1201 GTCTAGCCAAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGATGCTACGAG 1260

DB 1201 GTCTAGCCAAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGATGCTACGAG 1260

QY 1261 GATAAGGGTGTGCTATCCCCACCATATCGATCTTGGTGATTCGGCTGTGGTCATTCACT 1320

DB 1261 GATAAGGGTGTGCTATCCCCACCATATCGATCTTGGTGATTCGGCTGTGGTCATTCACT 1320

QY 1321 GATTATGACCAACAGCAGATGAGCAGGATCGGCCCAACCCGCTCGCTGCGCACTCTCGGCT 1380

DB 1321 GATTATGACCAACAGCAGATGAGCAGGATCGGCCCAACCCGCTCGCTGCGCACTCTCGGCT 1380

QY 1381 TTTTCTGTTCTCGAGCAAAATGATGTAATTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCT 1440

DB 1381 TTTTCTGTTCTCGAGCAAAATGATGTAATTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCT 1440

QY 1441 CAGTCCACTTACGGTCTGTCACCTGCGCGGTTTATATCTCGACAGAGCTGACTTTGGTG 1500

DB 1441 CAGTCCACTTACGGTCTGTCACCTGCGCGGTTTATATCTCGACAGAGCTGACTTTGGTG 1500

QY 1501 AATGTTGCACTGGCGCGCAGGATCGGCCCAACCCGCTCGCTGCGCACTCTCGGCT 1560

DB 1501 AATGTTGCACTGGCGCGCAGGATCGGCCCAACCCGCTCGCTGCGCACTCTCGGCT 1560

QY 1561 GACGGCGGCGGCTCTCGGCTGTCGAGCAATTTCCAGCAATTTCTTGTCTCTCTCTCTCTCT 1620

DB 1561 GACGGCGGCGGCTCTCGGCTGTCGAGCAATTTCCAGCAATTTCTTGTCTCTCTCTCTCTCT 1620

QY 1621 CGTGGCAAGCTCTCTCTTTTGGAGGCGCGCAACAAAGAGGATTTATCTTATATATAT 1680

DB 1621 CGTGGCAAGCTCTCTCTTTTGGAGGCGCGCAACAAAGAGGATTTATCTTATATATAT 1680

QY 1681 AATACTACTGCTAGTACAGATTTGATTTGAAATGTCGCGGCGCATCGGTCGCAAT 1740

DB 1681 AATACTACTGCTAGTACAGATTTGATTTGAAATGTCGCGGCGCATCGGTCGCAAT 1740

QY 1741 TCAACCTATACCAACAGGCTTGGGCGGCTCTCGGTCGCAATTTCTGCGCGCGGTTTG 1800

DB 1741 TCAACCTATACCAACAGGCTTGGGCGGCTCTCGGTCGCAATTTCTGCGCGCGGTTTG 1800

QY 1801 GCTCCACGCTCCGCGCTGCTCTGCTGGAGGATCTTTTGTATTCGCGGCGGCGCAC 1860

DB 1801 GCTCCACGCTCCGCGCTGCTCTGCTGGAGGATCTTTTGTATTCGCGGCGGCGCAC 1860

QY 1861 ACATTTGATGACTTCTGCGCTGAAATGCGCGCTTTTAGGCTCCAGGCTTGTCTTCCAG 1920

|||||

Db 1861 ACATTGTGATCTCTGCGCTGAATGCGCGCTTTAGGCTCCAGGGTTGTCTTCCAG 1920

QY 1921 TCAACTGTCGTGAGCTCCAGGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTGTAG 1980

Db 1921 TCAACTGTCGTGAGCTCCAGGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTGTAG 1980

RESULT 2

AAQ47130

ID AAQ47130 standard; DNA; 2100 BP.

XX AC AAQ47130;

XX AC

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 11-JAN-1994 (first entry)

XX DE

XX DE HEV third and second ORFs.

XX KW

XX KW Enterically transmitted non-A non-B hepatitis; ET-NANB; vaccine; ss.

XX OS

XX OS Hepatitis E virus; Mexico strain.

XX FH

XX FH Location/Qualifiers

FT CDS 5..376

FT FT /*tag= a

FT FT /label= ORF3

FT CDS 46..2100

FT FT /*tag= b

FT FT /label= ORF2

FT FT 275..376

FT FT /*tag= c

FT FT /label= 406.4-2

FT FT 718..2026

FT FT /*tag= d

FT FT /label= C2

FT FT 1045..2026

FT FT /*tag= e

FT FT /label= SG3

FT FT 1882..2026

FT FT /*tag= f

FT FT /label= 406.3-2

XX WO9314116-Al.

XX PN

XX PN 22-JUL-1993.

XX PD

XX PF 15-JAN-1993; 93WO-US000459.

XX PF 17-JAN-1992; 92US-00822335.

XX PR 01-MAY-1992; 92US-00876941.

XX PR

XX (GENE-) GENELABS TECHNOLOGIES INC.

XX PA (USSH) US SEC DEPT HEALTH.

XX XX

XX Reyes GR, Bradley DW, Tam AW, Carl M;

XX PI

XX WPI; 1993-243144/30.

XX DR

XX P-PSDB; AAR38786, AAR38787.

XX XX

XX New immunogenic hepatitis E virus (HEV) peptide(s) - are from the ORF1, ORF2 and ORF3 regions of HEV, useful as a vaccine against HEV infection.

XX XX

XX Disclosure; Fig 7; 48pp; English.

XX XX

XX Immunogenic hepatitis E virus (HEV) peptides are selected from the ORF1, ORF2 and ORF3 regions of HEV. The peptides can be used in vaccines to prevent infection by HEV. The antibodies can neutralise and block HEV infection and can be used to prevent or treat HEV infection. The peptides and antibodies can also be used as diagnostic reagents. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS field)

XX SQ Sequence 2100 BP; 381 A; 659 C; 516 G; 544 T; 0 U; 0 Other;

Query Match 100.0%; Score 1980; DB 2; Length 2100;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTAGGCTCTTTTGTCTGTTGTTCTCTCTGTTTCTGCTATGTTGCTGCGGCA 60

Db 46 ATGGCCCTAGGCTCTTTTGTCTGTTGTTCTCTCTGTTTCTGCTATGTTGCTGCGGCA 105

QY 61 CCGACCGGTTCAGCCGCTCTGGCGCGCGCTGTGGCGCGCGAGCGCGGTACCGCGCGGTGT 120

Db 106 CCGACCGGTTCAGCCGCTCTGGCGCGCGCTGTGGCGCGCGAGCGCGGTACCGCGCGGTGT 165

QY 121 TTCTGGGTGACCGGGTTGATTCTCAGCCCTTGGCAATCCCTATATTCATCAACCAAC 180

Db 166 TTCTGGGTGACCGGGTTGATTCTCAGCCCTTGGCAATCCCTATATTCATCAACCAAC 225

QY 181 CCCTTTGCCCGCAGAGTTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 240

Db 226 CCCTTTGCCCGCAGAGTTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 285

QY 241 CGGCACCTTGGCTCCACCTTGGCGAGATCAGGCGCCAGCGCCCTCCGCTGCGCTCCCGTGC 300

Db 286 CGGCACCTTGGCTCCACCTTGGCGAGATCAGGCGCCAGCGCCCTCCGCTGCGCTCCCGTGC 345

QY 301 CGACTGCCACAGCGGGGCTGCGGCGGTGACGGCTGTGGCGCGCTGCCCATGACACTCA 360

Db 346 CGACTGCCACAGCGGGGCTGCGGCGGTGACGGCTGTGGCGCGCTGCCCATGACACTCA 405

QY 361 CCGGTCCCGGACGTTGATTCTCGGGGTGCAATCTACGCGCGCGCTGCAATTTGTTGTTACT 420

Db 406 CCGGTCCCGGACGTTGATTCTCGGGGTGCAATCTACGCGCGCGCTGCAATTTGTTGTTACT 465

QY 421 TCACCCCTGACATCCTCTGTGGCGCTCTGGCACTAATTTAGTCTCTGTATGCGACCCCTTT 480

Db 466 TCACCCCTGACATCCTCTGTGGCGCTCTGGCACTAATTTAGTCTCTGTATGCGACCCCTTT 525

QY 481 AATCGCGCTCTGGCGCTGACGAGCGGTACTAATACTACATTATGGCCACAGAGCCCTCC 540

Db 526 AATCGCGCTCTGGCGCTGACGAGCGGTACTAATACTACATTATGGCCACAGAGCCCTCC 585

QY 541 AATTATGCACAGTACCGGGTTGCGCGCGCTACTATCCGTTACCGGCCCTTAGTGCTAAT 600

Db 586 AATTATGCACAGTACCGGGTTGCGCGCGCTACTATCCGTTACCGGCCCTTAGTGCTAAT 645

QY 601 GCAGTTGGAGGCTATGCTATATCCATTCTTTCTGGCGCTCAAAACCAACACCCCTACA 660

Db 646 GCAGTTGGAGGCTATGCTATATCCATTCTTTCTGGCGCTCAAAACCAACACCCCTACA 705

QY 661 TCTGTTGACATGAATTCATTACTTCCACTGATGTGAGGATTCCTGTTCAACCTGGCATA 720

Db 706 TCTGTTGACATGAATTCATTACTTCCACTGATGTGAGGATTCCTGTTCAACCTGGCATA 765

QY 721 GCATCTGAATTGTTTCATCCCAAGCGGCGCTTCACTACCGCAATCAAGTTGGCGCTCG 780

Db 766 GCATCTGAATTGTTTCATCCCAAGCGGCGCTTCACTACCGCAATCAAGTTGGCGCTCG 825

QY 781 GTTGAGACATCTGGTGTGTTGCTGAGGAGGAGCCACCTCCCGTCTTGTGATGATGATA 840

Db 826 GTTGAGACATCTGGTGTGTTGCTGAGGAGGAGCCACCTCCCGTCTTGTGATGATGATA 885

QY 841 CATGGCTCTCCAGTTAACTCCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 900

Db 886 CATGGCTCTCCAGTTAACTCCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 945

QY 901 GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTACACCTGTATACCAATACAGTGTG 960

Db 946 GACTTTGCCCTTAGAGCTTGAGTTTCGCAATCTACACCTGTATACCAATACAGTGTG 1005

QY 961 TCCGGTTACTCCAGCACTGCTCGTCACTCCGCGCGGCGCGGAGGAGTCCGAGCTG 1020

XX WPI; 1998-582599/49.
XX Hepatitis E virus proteins - useful for diagnosis or vaccine production
XX PT the virus.
XX PS
XX Claim 4; Col 69-76; 47pp; English.
XX The present sequence represents the composite Mexico strain sequence of
CC enterally transmitted non A non B hepatitis virus (ET-NANB) (hepatitis
CC E virus (HEV)). The specification describes an isolated protein which is
CC specifically immunoreactive with antibodies present in individuals
CC infected with HEV and encoded by a sequence contained in an open reading
CC frame (ORF) of an HEV genome. The genome has a sequence that is more than
CC 70% identical to the ORF1 sequence from Burma HEV isolate (AAV6321). The
CC protein is used as a vaccine and a diagnostic probe for ET-NANB. (Updated
CC on 25-MAR-2003 to correct PR field.) (Updated on 27-AUG-2003 to correct
CC OS field.)
XX
SQ Sequence 7171 BP; 1285 A; 2147 C; 1907 G; 1832 T; 0 U; 0 Other;
Query Match 100.0%; Score 1980; DB 2; Length 7171;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTAGGCTCTTTGGCTGTGTTCTCTCTGCTTCTGCTATGTTGCCGCGCCA 60
DB 5117 ATGCGCCCTAGGCTCTTTGGCTGTGTTCTCTCTGCTTCTGCTATGTTGCCGCGCCA 5176
QY 61 CCGACCGGTGACCGGTGATTTCTAGCCCTTCGGAATCCCTATATTCATCCAAACCAAC 180
DB 5177 CCGACCGGTGACCGGTGATTTCTAGCCCTTCGGAATCCCTATATTCATCCAAACCAAC 5236
QY 121 TTCTGGGGTGACCGGTGATTTCTAGCCCTTCGGAATCCCTATATTCATCCAAACCAAC 180
DB 5237 TTCTGGGGTGACCGGTGATTTCTAGCCCTTCGGAATCCCTATATTCATCCAAACCAAC 5296
QY 181 CCTTTGGCCCAAGACGTTGCGGTGCTGCGGTCTGGACCTCGCCCTTCGCCAACCAAGCC 240
DB 5297 CCTTTGGCCCAAGACGTTGCGGTGCTGCGGTCTGGACCTCGCCCTTCGCCAACCAAGCC 5356
QY 241 CCGCCACTTGGCTTCCACTTGGCGAGATCAGGCCAGCCGCTCGCTCGCTCGCTCGCTCGC 300
DB 5357 CCGCCACTTGGCTTCCACTTGGCGAGATCAGGCCAGCCGCTCGCTCGCTCGCTCGCTCGC 5416
QY 301 CGACCTGCCACAGCCGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 360
DB 5417 CGACCTGCCACAGCCGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 5476
QY 361 CCGCTCCCGGACGTTGATTTCTGCGGTGCAATTTCTAGCGCGCCAGTATATTTGTTCTACT 420
DB 5477 CCGCTCCCGGACGTTGATTTCTGCGGTGCAATTTCTAGCGCGCCAGTATATTTGTTCTACT 5536
QY 421 TCACCCCTGACATCTCTGTGGCTCTGGACCTAAATTTAGTCTGTATGAGCCGCTCGCT 480
DB 5537 TCACCCCTGACATCTCTGTGGCTCTGGACCTAAATTTAGTCTGTATGAGCCGCTCGCT 5596
QY 481 AATCCGCTCTGCGGCTGAGGACGCTACTAATCTCAATTAATGAGCCAGAGGCTCG 540
DB 5597 AATCCGCTCTGCGGCTGAGGACGCTACTAATCTCAATTAATGAGCCAGAGGCTCG 5656
QY 541 AATTAATGACAGTACCGGTTGCGGCTGCTACTATCCGTTTACCGGCTCTAGTGCCTAAT 600
DB 5657 AATTAATGACAGTACCGGTTGCGGCTGCTACTATCCGTTTACCGGCTCTAGTGCCTAAT 5716
QY 601 GCAGTTGGAGGCTATGCTATATCAATTTCTTCTGGCTCAAAACCAACCCCTACA 660
DB 5717 GCAGTTGGAGGCTATGCTATATCAATTTCTTCTGGCTCAAAACCAACCCCTACA 5776
QY 661 TCTGTTGACATGAATTCATTTCTTCTTCTGAGTGTGAGGATTTCTTCTCAACCTGGCATA 720
DB 5777 TCTGTTGACATGAATTCATTTCTTCTTCTGAGTGTGAGGATTTCTTCTCAACCTGGCATA 5836

QY 721 GCATCTGAATTGGTCTATCCCAAGCGAGCGCCTTCACTACCGCAATCAAGTTGGCGCTCG 780
DB 5837 GCATCTGAATTGGTCTATCCCAAGCGAGCGCCTTCACTACCGCAATCAAGTTGGCGCTCG 5896
QY 781 GTTGAGACATCTGGTGTGTTCTGAGGAGGAAGCACTCCGGTCTTGTCATGTTATCAATA 840
DB 5897 GTTGAGACATCTGGTGTGTTCTGAGGAGGAAGCACTCCGGTCTTGTCATGTTATGCATA 5956
QY 841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 900
DB 5957 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 6016
QY 901 GACTTTGCTCTAGAGCTTGAGTTTGCATCTCACCACTCTAAACCAATACACAGTGTG 960
DB 6017 GACTTTGCTCTAGAGCTTGAGTTTGCATCTCACCACTCTAAACCAATACACAGTGTG 6076
QY 961 TCCCGTTACTCCAGCACTGCTCTGCTCACTCCGCGGAGGCGGACGCGGACTGCGGAGCTG 1020
DB 6077 TCCCGTTACTCCAGCACTGCTCTGCTCACTCCGCGGAGGCGGACGCGGACTGCGGAGCTG 6136
QY 1021 ACCCAACTGCGAGCCACCAAGTTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080
DB 6137 ACCCAACTGCGAGCCACCAAGTTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA 6196
QY 1081 GGTGAAGTCCGCGCGGAGTAGCTCTAACATTAATTAACCTTGTCTGACACGCTCTCTCGGC 1140
DB 6197 GGTGAAGTCCGCGCGGAGTAGCTCTAACATTAATTAACCTTGTCTGACACGCTCTCTCGGC 6256
QY 1141 GGGTCCCGACAGAAATTAATTTGCTGCGTGGGGGCAACTGTTTATTTCCGCGCCGCTT 1200
DB 6257 GGGTCCCGACAGAAATTAATTTGCTGCGTGGGGGCAACTGTTTATTTCCGCGCCGCTT 6316
QY 1201 GTCTCAGCAATGCGGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAAATCTCAGCAG 1260
DB 6317 GTCTCAGCAATGCGGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAAATCTCAGCAG 6376
QY 1261 GATAAGGGTGTGCTATCCCGACAGATATCGATCTTGTGTGATTTGGGTGATTTGCTCAG 1320
DB 6377 GATAAGGGTGTGCTATCCCGACAGATATCGATCTTGTGTGATTTGGGTGATTTGCTCAG 6436
QY 1321 GATTATGACAAACAGCATGAGCAGAGATCGGCCCAACCGCTGCGCTCGCTCGGCACTCGGCCT 1380
DB 6437 GATTATGACAAACAGCATGAGCAGAGATCGGCCCAACCGCTGCGCTCGCTCGGCACTCGGCCT 6496
QY 1381 TTTTCTGTTCTCGGCAATGATGTTACTTTGGCTCTCTCTCCTCCTCAGCCGAGATGAC 1440
DB 6497 TTTTCTGTTCTCGGCAATGATGTTACTTTGGCTCTCTCTCCTCCTCAGCCGAGATGAC 6556
QY 1441 CAGTCCACTTACGGGTGCTCAACTGGCCGCGTTTATATCTCGGACAGCGTGAATTTGGTG 1500
DB 6557 CAGTCCACTTACGGGTGCTCAACTGGCCGCGTTTATATCTCGGACAGCGTGAATTTGGTG 6616
QY 1501 AATGTTGCACTGCGCGCAGCGCTGAGCCGATCGCTTGAATGCTGCTCAAGTCAACCTC 1560
DB 6617 AATGTTGCACTGCGCGCAGCGCTGAGCCGATCGCTTGAATGCTGCTCAAGTCAACCTC 6676
QY 1561 GACGCGCGCCCTCCGACTGTTGAGCAATATTCAGACATTTCTTGTGCTCCCGCTT 1620
DB 6677 GACGCGCGCCCTCCGACTGTTGAGCAATATTCAGACATTTCTTGTGCTCCCGCTT 6736
QY 1621 CGTGCAAGCTCTCTCTTTTGGAGCGCGGCAACAAAAGCAGGTTTATCTTATAATAT 1680
DB 6737 CGTGCAAGCTCTCTCTTTTGGAGCGCGGCAACAAAAGCAGGTTTATCTTATAATAT 6796
QY 1681 AATACTACTGCTAGTACAGATTTCTGATTGAATGCTGCGGCGCATCGGGTCGCATT 1740
DB 6797 AATACTACTGCTAGTACAGATTTCTGATTGAATGCTGCGGCGCATCGGGTCGCATT 6856
QY 1741 TCAACCTATACCAACAGGCTTGGGGCGGTCGCGTCCGCTCGCATTTCTCGCGCGCGGTTTG 1800
DB 6857 TCAACCTATACCAACAGGCTTGGGGCGGTCGCGTCCGCTCGCATTTCTCGCGCGCGGTTTG 6916
QY 1801 GCTCCACGCTCCGCGCTCGCTCTGCTGAGGATACATTTTGAATTTATCGGGGCGGCGCAC 1860

Qy	1201	GTCTACGCCAATGGCGAGCCAAACCGTGAAGCTCTTATACATCAGTGGAGAAATGCTACGAG	1260
Db	6317	GTCTACGCCAATGGCGAGCCAAACCGTGAAGCTCTTATACATCAGTGGAGAAATGCTACGAG	6376
Qy	1261	GATAAGGGTGTGGTATCTCCCGACGATATCGATCTTTGGTGATTCGGGTGCGGTCAATTCAG	1320
Db	6377	GATAAGGGTGTGGTATCTCCCGACGATATCGATCTTTGGTGATTCGGGTGCGGTCAATTCAG	6436
Qy	1321	GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCCCGCTGCCTCGGCCCATCTCGGCCT	1380
Db	6437	GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCCCGCTGCCTCGGCCCATCTCGGCCT	6496
Qy	1381	TTTTCTGTTCTCCGAGCAAAATGATGTACTTTTGGGTGTCCTCACTCGAGCCGAGTATGAC	1440
Db	6497	TTTTCTGTTCTCCGAGCAAAATGATGTACTTTTGGGTGTCCTCACTCGAGCCGAGTATGAC	6556
Qy	1441	CAGTCCACTTACGGGTGCTCAACTGGGCCCGGTTTATATCTCGACACAGCTGACTTTGGTG	1500
Db	6557	CAGTCCACTTACGGGTGCTCAACTGGGCCCGGTTTATATCTCGACACAGCTGACTTTGGTG	6616
Qy	1501	AATGTTCCGACTGGCGCGCAGGCGGTAGCCCGATCGCTTGACCTGGTCCAAAGTCACCCCTC	1560
Db	6617	AATGTTCCGACTGGCGCGCAGGCGGTAGCCCGATCGCTTGACCTGGTCCAAAGTCACCCCTC	6676
Qy	1561	GACGGCGGCCCTCCGACACTGTTGAGCAATATTTCCAAGACATTTCTTGTGCTCCGCCCTT	1620
Db	6677	GACGGCGGCCCTCCGACACTGTTGAGCAATATTTCCAAGACATTTCTTGTGCTCCGCCCTT	6736
Qy	1621	CGTGGCAAGCTCTCCTTTTGGGAGCCCGGCACACAAAGCAGGTTATCCTTATAATTAT	1680
Db	6737	CGTGGCAAGCTCTCCTTTTGGGAGCCCGGCACACAAAGCAGGTTATCCTTATAATTAT	6796
Qy	1681	AATACTACTGCTAGTGACAGATCTCATTTGAAATGCTCGCGGCCATCGGTCGCCATT	1740
Db	6797	AATACTACTGCTAGTGACAGATCTCATTTGAAATGCTCGCGGCCATCGGTCGCCATT	6856
Qy	1741	TCAACTATACCAACAGGCTTGGGGCCGGTCCGGTCGCCATTTCTCGGCCCGCGGTTTG	1800
Db	6857	TCAACTATACCAACAGGCTTGGGGCCGGTCCGGTCGCCATTTCTCGGCCCGCGGTTTG	6916
Qy	1801	GCTCCAGCTCCGGCCTTGGCTCTGCTGGAGGATCTTTTGATTTATCCGGGGCGGGCGCAC	1860
Db	6917	GCTCCAGCTCCGGCCTTGGCTCTGCTGGAGGATCTTTTGATTTATCCGGGGCGGGCGCAC	6976
Qy	1861	ACATTTGATGACTTCTGCGCTGAATCGCGCTTTTAGGCTTCCAGGGTGTGCTTTCAG	1920
Db	6977	ACATTTGATGACTTCTGCGCTGAATCGCGCTTTTAGGCTTCCAGGGTGTGCTTTCAG	7036
Qy	1921	TCAACTGTGCTGAGCTCCAGCGCTTTAAAGTTAAGGTGGGTAAAACCTCGGAGTGTGTAG	1980
Db	7037	TCAACTGTGCTGAGCTCCAGCGCTTTAAAGTTAAGGTGGGTAAAACCTCGGAGTGTGTAG	7096

XX	07-JUN-1995;	95US-00478507.
PF		
XX	17-JUN-1988;	88US-00208997.
PR	11-APR-1989;	89US-00336672.
PR	16-JUN-1989;	89US-00367486.
PR	13-OCT-1989;	89US-00420921.
PR	05-APR-1990;	90US-00505888.
PR	05-APR-1991;	91US-00681078.
PR	25-JUL-1994;	94US-00279823.
XX	(GENE-) GENELABS TECHNOLOGIES INC.	
PA	(USGO) US GOVERNMENT.	
XX	Yarborough PO, Krawczynski KZ, Fry KE, Bradley DW, Tam A, Reyes GR;	
XX	WPI; 2000-593712/56.	
XX	Identifying recombinant antigen immunoreactive with antibody induced by	
PT	hepatitis E virus (HEV), for detecting HEV infection, comprises	
PT	immunoreacting a polypeptide from an HEV genome with an HEV-positive	
PT	antiseraum.	
XX		
XX	Claim 5; Col 69-76; 46pp; English.	
XX	The present invention describes a method for identifying a recombinant	
CC	antigen immunoreactive with a hepatitis E virus (HEV)-induced antibody.	
CC	The method comprises producing a polypeptide derived from an HEV genome,	
CC	immunoreacting the polypeptide with an HEV-positive antiseraum and	
CC	selecting the polypeptide as a recombinant antigen if the polypeptide	
CC	reacts with the HEV-positive antiseraum. The method is useful for	
CC	identifying recombinant antigen immunoreactive with antibody induced by	
CC	HEV. The enterically transmitted non-A/non-B hepatitis (ET-NANB) (also	
CC	known as HEV)-specific fragments are useful for identifying ET-NANB-	
CC	derived cDNAs, which contain additional sequence information, as primers	
CC	for detecting ET-NANB viral genomic material in a patient sample, for the	
CC	synthesis of polypeptides for use in immunoassays, and for identifying	
CC	similar antigenic regions encoded by related viral strains, e.g. Burmese	
CC	strain. The antigens are especially useful in the preparation of vaccine	
CC	against ET-NANB infection. These antigens may further be used to prepare	
CC	antibodies to ET-NANB virus particles for use directly as antiviral	
CC	agents, and to produce antiseraum designed for pre- or post-exposure	
CC	prophylaxis. The present sequence represents a specifically claimed HEV	
CC	composite Mexican strain gene sequence for use in the present invention	
XX		
SQ	Sequence 7171 BP; 1285 A; 2147 C; 1907 G; 1832 T; 0 U; 0 Other;	
Query Match 100.0%; Score 1980; DB 3; Length 7171;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 ATGGCGCCTAGGCGCTCTTTTGCTGTTCCTCTTGTTCCTATGTCCTATGTCGCCGCCA 60	
Dd	5117 ATGGCGCCTAGGCGCTCTTTTGCTGTTCCTCTTGTTCCTATGTCGCCGCCA 5176	
Qy	61 CCGACCGGTGACCGGTCTGGCGCGCTGCTGGCGCGCGCGCGCGCGTACGGCGGTGGT 120	
Dd	5177 CCGACCGGTGACCGGTCTGGCGCGCTGCTGGCGCGCGCGCGCGTACGGCGGTGGT 5236	
Qy	121 TTCTGGGGTGACCGGGTGATTCTCAGCCCTTCGCATCCCTATATTCATCAACAAC 180	
Dd	5237 TTCTGGGGTGACCGGGTGATTCTCAGCCCTTCGCATCCCTATATTCATCAACAAC 5296	
Qy	181 CCTTTTCCCCAGAGTTGCGCTGCGTGGCGGTGTGAACCTCGCCTTGCCCAACAGCC 240	
Dd	5297 CCTTTTCCCCAGAGTTGCGCTGCGTGGCGGTGTGAACCTCGCCTTGCCCAACAGCC 5356	
Qy	241 CGGCACTTGGCTCCACTTCGGCGAGATCAGGCCAGCGCCCTCCGCTGCTCCGCTCGC 300	
Dd	5357 CGGCACTTGGCTCCACTTCGGCGAGATCAGGCCAGCGCCCTCCGCTGCTCCGCTCGC 5416	
Qy	301 CGACTTGCCACAGCGGGGCTGGCGGCGCTGACCGCTGTGGCGCTGCCCATGACACTCA 360	

[illegible]

QY 1141 GGGCTCCGACAGAAATTAATTTTCGTGGTGGCGGCAACTGTTTATTTCCCGCCGGTT 1200
DB |||||
QY 1201 GTCTCAGCCAAATGGCGAGCAACCGGTGAAGCTCTATACATCAGTGGAGAATGCTCAGCAG 1260
DB |||||
QY 1261 GATAAGGGTGTGCTATCCCCACGATATCGATCTTTGGTGTATCGGTGTGCTCATTTCAG 1320
DB |||||
QY 1321 GATTATGACAAACAGATGAGCAGGATCGSCCAACCCCGTCGCTCGCCATCTCCGCGCT 1380
DB |||||
QY 1381 TTTTCTGTCTCCGAGCAATGATGATCTTTGGCTGTCCCTCACTCAGCCGAGTATGAC 1440
DB |||||
QY 1441 CAGTCCACTACGGGTGCTCACTGGCCCGGTTTATCTCGACAGCGTGAATTTGGTG 1500
DB |||||
QY 1501 AATGTTGCGACTCGCGCAGGCGGTAGCCCGATCGCTTGACTGGTCCAAAGTCAACCTC 1560
DB |||||
QY 1561 GACGGCGGCGCTCCGACTGTGAGCAATATTCGAAGCATCTTTGTGCTCCCGCTT 1620
DB |||||
QY 1621 CGTGGCAAGCTCTCTTTGGGAGCGCGCACACAAAGCAGSTTATCCTTATAATTAT 1680
DB |||||
QY 1681 AATACTACTGCTAGTGACAGATTCTGATTGAAAATGCTCGCGGCCATCGGGTCGCCATT 1740
DB |||||
QY 1741 TCAACCTATACCAACAGCGTGGGCGGTGCGGTGCGCATTTCTGCGGCGCGGTTTG 1800
DB |||||
QY 1801 GCTCCAGCTCGCCCTGCTGCTGGAGGATACCTTTGATTATCGGGGCGGCGCAC 1860
DB |||||
QY 1861 ACATTTGATGACTTCTGCCCTGAATGCGCGCTTTAGGCCCTCCAGGGTGTGCTTTCCAG 1920
DB |||||
QY 1921 TCAACTGCTGAGCTCCAGCGCTTAAAGTTAAAGTGGGTAAACTCGGGAGTTGTAG 1980
DB |||||
QY 7037 TCAACTGCTGAGCTCCAGCGCTTAAAGTTAAAGTGGGTAAACTCGGGAGTTGTAG 7096

RESULT 9

ADD24378 standard; cDNA; 7171 BP.

ID ADD24378

XX AC ADD24378;

XX DT 15-JAN-2004 (first entry)

XX DE Hepatitis E virus strain Mexico composite cDNA #1.

XX KW ss; gene; non-A/non-B viral hepatitis; ET-NANB; HEV; vaccine.

XX OS Hepatitis E virus; strain Mexico.

XX XX US2003124510-A1.

PD 03-JUL-2003.
XX 07-MAY-2001; 2001US-00851410.
XX 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-00336672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 05-APR-1991; 91US-00681078.
PR 25-JUL-1994; 94US-00279823.
PR 03-AUG-1998; 98US-00128275.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam AW;
PI Fry KE;
XX WPI; 2003-810978/76.
XX New proteins derived from an enterically transmitted non-A/non-B (ET-
PT NANB) viral hepatitis agent, useful for diagnosing ET-NANB infection and
PT producing vaccines for immunizing an individual against ET-NANB hepatitis
PT viral agent.
XX Claim 3; SEQ ID NO 10; 81pp; English.
XX The invention relates to a protein derived from an enterically
CC transmitted non-A/non-B viral hepatitis agent. The viral proteins are
CC useful for detecting infection by the enterically transmitted non-A/non-B
CC (ET-NANB/HEV) hepatitis viral agent. The proteins are also useful for
CC producing vaccines for immunising an individual against HEV hepatitis
CC viral agent. The DNA fragments are useful for identifying and sequencing
CC the entire viral agent and for assaying the viral agent in an infected
CC sample. The present sequence is used in the exemplification of the
CC invention.

SQ Sequence 7171 BP; 1285 A; 2147 C; 1907 G; 1832 T; 0 U; 0 Other;

Query Match 100.0%; Score 1980; DB 9; Length 7171;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGCGCCCTAGGCGCTCTTTTGTGCTGTTTCTCTTGTCTGCTATGTTGCCGCGCCA 60
DB 5117 ATCGCGCCCTAGGCGCTCTTTTGTGCTGTTTCTCTTGTCTGCTATGTTGCCGCGCCA 5176
QY 61 CCGACCGGTGAGCGGTCTGCGCCCGCTGCTGGGCGGCGAGCGGTACCGGGGTGT 120
DB 5177 CCGACCGGTGAGCGGTCTGCGCCCGCTGCTGGGCGGCGAGCGGTACCGGGGTGT 5236
QY 121 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
DB 5237 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 5296
QY 181 CCCTTTTGCCCCAGACGTTGCGCGTCTGCGGTCTGGAACCTCGCTTCGCCAACACGAC 240
DB 5297 CCCTTTTGCCCCAGACGTTGCGCGTCTGCGGTCTGGAACCTCGCTTCGCCAACACGAC 5356
QY 241 CGGCGACTTGGCTCCACTTTGGCGAGATCAGGCCCGCGCCCTCCGCTGCCCTCCGTCG 300
DB 5357 CGGCGACTTGGCTCCACTTTGGCGAGATCAGGCCCGCGCCCTCCGCTGCCCTCCGTCG 5416
QY 301 CGACCTGCCACAGCCGGGGCTGCGGCGCTGACGCGCTGTGGCGGTGCCCATGACACTCA 360
DB 5417 CGACCTGCCACAGCCGGGGCTGCGGCGCTGACGCGCTGTGGCGGTGCCCATGACACTCA 5476
QY 361 CCGCTCCCGGACGTTGATTCTCGCGGTGCAATTTCTACCGCGCGCATTAATTTGCTACT 420
DB 5477 CCGCTCCCGGACGTTGATTCTCGCGGTGCAATTTCTACCGCGCGCATTAATTTGCTACT 5536
QY 421 TCACCCCTGACATCTCTGTGGCTCTGCGCACTAAATTTAGTCTGTATGACAGCCCCCTT 480

Query Match									
Best Local Similarity 100.0%; Score 1980; DB 9; Length 7180;									
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
1	ATGCGCCCTAGGCGCTCTTTTGGCTGTGTTCTCTTTGTTCTGCTATGTTGGCCCGGCCA	60							
5117	ATGCGCCCTAGGCGCTCTTTTGGCTGTGTTCTCTTTGTTCTGCTATGTTGGCCCGGCCA	5176							
61	CGACCGGTGACCGGTCTGGCGCGCGCTGTGGCGGGCGGACGGCGGTACCGGGGTGGT	120							
5177	CGACCGGTGACCGGTCTGGCGCGCGCTGTGGCGGGCGGACGGCGGTACCGGGGTGGT	5236							
121	TTCTGGGGTGACCGGGTGTATTTCTCAGCGCTTTGCGCAATCCCTATATTTCTATCCAAACCAAC	180							
5237	TTCTGGGGTGACCGGGTGTATTTCTCAGCGCTTTGCGCAATCCCTATATTTCTATCCAAACCAAC	5296							
181	CCCTTTGCCCGCAGAGTTGCCGCTGCGTCTCGGGTCTGGAAGTCTGCCCTTCGCCAACACAGCC	240							
5297	CCCTTTGCCCGCAGAGTTGCCGCTGCGTCTCGGGTCTGGAAGTCTGCCCTTCGCCAACACAGCC	5356							
241	CGGCCACTTGGCTTCACCTTGGGGAGATCAGGCCCGACGGCGCCCTCCGCTGCTCCGTCGC	300							
5357	CGGCCACTTGGCTTCACCTTGGGGAGATCAGGCCCGACGGCGCCCTCCGCTGCTCCGTCGC	5416							
301	CGACCTGCCACAGCCGGGGCTGGCGGCTGACGGCTGTGGCGCTGCCATGACACCTCA	360							
5417	CGACCTGCCACAGCCGGGGCTGGCGGCTGACGGCTGTGGCGCTGCCATGACACCTCA	5476							
361	CCCGTCCCGGAGTTGATTTCTCGCGGTGCAATTTCTACCGGCCAGTATAATTTGTCTACT	420							
5477	CCCGTCCCGGAGTTGATTTCTCGCGGTGCAATTTCTACCGGCCAGTATAATTTGTCTACT	5536							
421	TCACCCCTGACATCTCTGTGGCTCTGGGCACTAATTTAGTCTGTATGACAGCCGCCCTT	480							
5537	TCACCCCTGACATCTCTGTGGCTCTGGGCACTAATTTAGTCTGTATGACAGCCGCCCTT	5596							
481	AATCCGCTCTGCGGCTGCGAGCGGTACTAATCTACATATATGGCCACAGAGGCGCTCC	540							
5597	AATCCGCTCTGCGGCTGCGAGCGGTACTAATCTACATATATGGCCACAGAGGCGCTCC	5856							
541	AATATGACAGTACCGGGTGGCCGCTACTATCCGTTACCGGCCCTAGTGCCTAAT	600							
5657	AATATGACAGTACCGGGTGGCCGCTACTATCCGTTACCGGCCCTAGTGCCTAAT	5716							
601	GCAGTTGAGGCTATGCTATATTCATTTCTTCTGGGCTTAAACCAACACCCCTTACA	660							
5717	GCAGTTGAGGCTATGCTATATTCATTTCTTCTGGGCTTAAACCAACACCCCTTACA	5776							
661	TCGTGTGACATCAATTTCAATTTCTTCTGAGTGTGAGGATTTCTTCTCAACCTGGCATA	720							
5777	TCGTGTGACATCAATTTCAATTTCTTCTGAGTGTGAGGATTTCTTCTCAACCTGGCATA	5836							
721	GCATCTGAATTCGTTCATCCCAAGCAGCGGCTTCACTACCGCAATCAAGGTTGGCGCTG	780							
5837	GCATCTGAATTCGTTCATCCCAAGCAGCGGCTTCACTACCGCAATCAAGGTTGGCGCTG	5896							
781	GTGAGACATCTGGTGTGCTGAGGAGGACACCTCCGCTTGTGTCATGTTATGACATA	840							
5897	GTGAGACATCTGGTGTGCTGAGGAGGAGACACCTCCGCTTGTGTCATGTTATGACATA	5956							
841	CATGGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCGCTTACTG	900							
5957	CATGGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCGCTTACTG	6016							
901	GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACCACTGTAAACCAATACACGTGTG	960							
6017	GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACCACTGTAAACCAATACACGTGTG	6076							
961	TCCGTTACTCCAGCACTGCTGTCATCTCCGCGCGAGGGCGGACGGGACTGCGGAGCTG	1020							
6077	TCCGTTACTCCAGCACTGCTGTCATCTCCGCGCGAGGGCGGACGGGACTGCGGAGCTG	6136							

QY	1021	ACCACAACTGACGCCACAGGTTTCATGAAAGATCTCCACTTACCGGCCCTTAATGGGGTA	108
Db	6137	ACCACAACTGACGCCACAGGTTTCATGAAAGATCTCCACTTACCGGCCCTTAATGGGGTA	6196
QY	1081	GGTGAAGTTCGGCCGGGATAGCTCTAACATTACTTAACCTTGTCTGACAGCTCCTCGGC	1140
Db	6197	GGTGAAGTTCGGCCGGGATAGCTCTAACATTACTTAACCTTGTCTGACAGCTCCTCGGC	6256
QY	1141	GGGCTCCCGACAGAAATTAATTTCTGCGTCTGGCGGCAACTGTGTTTATTCGCCGCCGGTT	1200
Db	6257	GGGCTCCCGACAGAAATTAATTTCTGCGTCTGGCGGCAACTGTGTTTATTCGCCGCCGGTT	6316
QY	1201	GTCTCAGCAATGGCGAGCCAAACCGTGAAGCTCTATATCATCAGTGGAGAAATGCTCAGAG	1260
Db	6317	GTCTCAGCAATGGCGAGCCAAACCGTGAAGCTCTATATCATCAGTGGAGAAATGCTCAGAG	6376
QY	1261	GATAGGGTGTGCTATCCCCACGATATCGATCTTGGTATTCGCTGTGGTCAATTCAG	1320
Db	6377	GATAGGGTGTGCTATCCCCACGATATCGATCTTGGTATTCGCTGTGGTCAATTCAG	6436
QY	1321	GATTATGACAAACGACATGACAGGATCGGCCCAACCCCGTCCGCTGGCCCATCTCGGCCCT	1380
Db	6437	GATTATGACAAACGACATGACAGGATCGGCCCAACCCCGTCCGCTGGCCCATCTCGGCCCT	6496
QY	1381	TTTTCTGTTCTCCGAGCAAAATGATGTAATTTGGTGTCCCTCACTCAGCGCGAGTATGAC	1440
Db	6497	TTTTCTGTTCTCCGAGCAAAATGATGTAATTTGGTGTCCCTCACTCAGCGCGAGTATGAC	6556
QY	1441	CAGTCCACTTACGGGTCTGTAATGGCCCCGGTTTATATCTCGACAGGGTACTTTGGTG	1500
Db	6557	CAGTCCACTTACGGGTCTGTAATGGCCCCGGTTTATATCTCGACAGCGTACTTTGGTG	6616
QY	1501	AATGTTGGACTTGGCGCGCAGCGCTAGCCCCGATCGTTGACTGCTCCAAAGTCACCCCTC	1560
Db	6617	AATGTTGGACTTGGCGCGCAGCGCTAGCCCCGATCGTTGACTGCTCCAAAGTCACCCCTC	6676
QY	1561	GACGGCGGCGCCCTCCCGACTGTTGAGCAATATTTCCAAAGCAATTTCTTGTGCTCCCGCTT	1620
Db	6677	GACGGCGGCGCCCTCCCGACTGTTGAGCAATATTTCCAAAGCAATTTCTTGTGCTCCCGCTT	6736
QY	1621	CGTGGCAAGCTCTCTTTTGGAGCGCGGCAACAACAAAGCAGGTTATCCTTATAATAT	1680
Db	6737	CGTGGCAAGCTCTCTCTTTTGGAGCGCGGCAACAACAAAGCAGGTTATCCTTATAATAT	6796
QY	1681	AATACTACTGCTPAGTGACAGATTCCTGATTGAAATGCTGCCGCCATTCGGGTGCGCCATT	1740
Db	6797	AATACTACTGCTPAGTGACAGATTCCTGATTGAAATGCTGCCGCCATTCGGGTGCGCCATT	6856
QY	1741	TCAACTATACCAACAGGCTTTGGGCGGCTCCGCTCGCCATTTCTGCGGCGCGGGTTTG	1800
Db	6857	TCAACTATACCAACAGGCTTTGGGCGGCTCCGCTCGCCATTTCTGCGGCGCGGGTTTG	6916
QY	1801	GCTCCACGCTCCGCGCTGCTCTGCTGGAGATACCTTTTGATATTCGGGGCGGCGCAC	1860
Db	6917	GCTCCACGCTCCGCGCTGCTCTGCTGGAGATACCTTTTGATATTCGGGGCGGCGCAC	6976
QY	1861	ACATTTTGATGACTCTGCGCTGGAATGCCCGCTTTAGGCCCTCCAGGGTTGTGCTTCCAG	1920
Db	6977	ACATTTTGATGACTCTGCGCTGGAATGCCCGCTTTAGGCCCTCCAGGGTTGTGCTTCCAG	7036
QY	1921	TCAACTGTGCTGAGCTCCAGCGCCTTAAAGTTAAGTGGGTAAACCTCGGAGATTGTAG	1980
Db	7037	TCAACTGTGCTGAGCTCCAGCGCCTTAAAGTTAAGTGGGTAAACCTCGGAGATTGTAG	7096
RESULT 11			
ID	AAQ46814		
XX	AAQ46814 standard; DNA; 2100 BP.		
AC	AAQ46814;		
XX			
DT	25-MAR-2003 (revised)		
DT	14-FEB-1994 (first entry)		

Db	781	GTTGAGACATCTGGTGTTCGTGAGGAGGAAGCACCCTCCGGTCTTGTCTATGTTATGATGATA	840
Qy	841	CATGGCTCTCCAGTTAACTCCTATACCAATAACCCCTTTATACCGGTGCCTTGGCTTACTG	900
Db	841	CATGGCTCTCCAGTTAACTCCTATACCAATAACCCCTTTATACCGGTGCCTTGGCTTACTG	900
Qy	901	GACTTTGGCTTAGAGCTTGAGTTTCGCAATCTCAACCTGTATAACCAATAACACGTTG	960
Db	901	GACTTTGGCTTAGAGCTTGAGTTTCGCAATCTCAACCTGTATAACCAATAACACGTTG	960
Qy	961	TCCGGTTACTCCAGCACTGCTCGTCACTCCGCGCGAGGGCCGACGGAAGTGGGAG	1017
Db	961	TCCGGTTACTCCAGCACTGCTCGTCACTCCGCGCGAGGGCCGACGGAAGTGGGAG	1020
Qy	1018	CTGACCACAACTGCAGCGCACAGAGTTTCATGAAGATCTCCACTTACCGGCCCTTAATGGG	1077
Db	1021	CTGACCACAACTGCAGCGCACAGAGTTTCATGAAGATCTCCACTTACCGGCCCTTAATGGG	1080
Qy	1078	GTAGGTGAAGTGGCGCGGGGATAGTCTTAACATTAATTAACTTGTGTGACACGCTCCTC	1137
Db	1081	GTAGGTGAAGTGGCGCGGGGATAGTCTTAACATTAATTAACTTGTGTGACACGCTCCTC	1140
Qy	1138	GGCGGGCTCCCGACAGAAATTAATTTTCGTGGCTGGCGGCAACTGTGTTTATTTCCGCCCG	1197
Db	1141	GGCGGGCTCCCGACAGAAATTAATTTTCGTGGCTGGCGGCAACTGTGTTTATTTCCGCCCG	1200
Qy	1198	GTTGTCTCAGCAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAG	1257
Db	1201	GTTGTCTCAGCAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAG	1260
Qy	1258	CAGATTAAGGGTGTGTGATCCGCCACGATATCGATCTGGTGATTCGGGTGGTCAATT	1317
Db	1261	CAGATTAAGGGTGTGTGATCCGCCACGATATCGATCTGGTGATTCGGGTGGTCAATT	1320
Qy	1318	CAGATTAACAGCAACAGCATGACGAGATCGGCCCAACCCGCTGCCTCGCCCATCTCGG	1377
Db	1321	CAGATTAACAGCAACAGCATGACGAGATCGGCCCAACCCGCTGCCTCGCCCATCTCGG	1380
Qy	1378	CCTTTTTCTGTCTCCGAGCAATGATGTACTTTGGCTGTCCCTCACTGCAGCCGAGTAT	1437
Db	1381	CCTTTTTCTGTCTCCGAGCAATGATGTACTTTGGCTGTCCCTCACTGCAGCCGAGTAT	1440
Qy	1438	GACAGTCCACTTACGGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCGTGAATTG	1497
Db	1441	GACAGTCCACTTACGGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCGTGAATTG	1500
Qy	1498	GTAAATGTTGCGACTGCGCGCGAGCGGTAGCCGATCGCTTGACTGTGCCAAAGTCACC	1557
Db	1501	GTAAATGTTGCGACTGCGCGCGAGCGGTAGCCGATCGCTTGACTGTGCCAAAGTCACC	1560
Qy	1558	CTCGACGGCGGGCCCTCCGACTGTTTGAGCAATATTTCAAAGACATTTCTTTGTGCTCCCC	1617
Db	1561	CTCGACGGCGGGCCCTCCGACTGTTTGAGCAATATTTCAAAGACATTTCTTTGTGCTCCCC	1620
Qy	1618	CTTCGTGGCAAGCTCTCTTTTGGAGGGCGGCACAAACAAAGCAGGTTATCCTTATAAT	1677
Db	1621	CTTCGTGGCAAGCTCTCTTTTGGAGGGCGGCACAAACAAAGCAGGTTATCCTTATAAT	1680
Qy	1678	TATAATACTACTGTAGTAGACAGATTTGATTTGAAAATGCTGCGGGCCATCGGTCGCC	1737
Db	1681	TATAATACTACTGTAGTAGACAGATTTGATTTGAAAATGCTGCGGGCCATCGGTCGCC	1740
Qy	1738	ATTTCAACCTATACCAACAGCTTTGGGGCGGTCCTCGCTGCGCAATTTCTCGGCCCGGTT	1797
Db	1741	ATTTCAACCTATACCAACAGCTTTGGGGCGGTCCTCGCTGCGCAATTTCTCGGCCCGGTT	1800
Qy	1798	TTGGCTCCACGCTCCGCGCTGTCTGTGGAGGATACCTTTTGATTTATCCGGCGCGCGG	1857
Db	1801	TTGGCTCCACGCTCCGCGCTGTCTGTGGAGGATACCTTTTGATTTATCCGGCGCGCGG	1860
Qy	1858	CACACATTTGATGACTTTCGCCCTGAAATGCCGCGCTTTTAGCGCTCCAGGGTTGTGCTTTC	1917

Db	1861	CACACATTTGATGACTTCTGCGCCTGAATACGCGCGCTTTAGGCGCTCCAGGGTTGTGCTTTC	1920
Qy	1918	CAGTCAACTGTCGCTGAGCTCCAGCGCCTTAAAGCTTAAAGTGGGTTAAAGTTCGGGAGTTG	1977
Db	1921	CAGTCAACTGTCGCTGAGCTCCAGCGCCTTAAAGTGGGTTAAAGTTCGGGAGTTG	1980
Qy	1978	TAG 1980	
Db	1981	TAG 1983	
RESULT 13			
AAT27110			
ID	AAT27110	standard; DNA; 1647 BP.	
XX	AC	AAT27110;	
XX	DT	16-OCT-2003 (revised)	
DT	06-AUG-1996	(first entry)	
XX	DE	Hepatitis E virus (Mexico strain) 62K antigen DNA.	
XX	KW	HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;	
KW	diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;		
KW	capsid; ds.		
XX	OS	Hepatitis E virus; Mexico strain.	
XX	PN	W09612807-A2.	
XX	PD	02-MAY-1996.	
XX	PF	23-OCT-1995; 95WO-US013703.	
XX	PR	24-OCT-1994; 94US-00327952.	
PR	13-OCT-1995; 95US-00542634.		
XX	PA	(GENE-) GENELABS TECHNOLOGIES INC.	
XX	PI	Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;	
XX	DR	WPI; 1996-230608/23.	
DR	DR	P-PSDB; AAR96092.	
XX	PT	Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as	
PT	PT	diagnostic reagents for determining HEV infection and in vaccines.	
XX	PS	Disclosure; Page 73-74; 125pp; English.	
XX	CC	A DNA sequence (AAT27110) codes for 62K antigen (AAR96092) comprising the	
CC	CC	C-terminal 549 amino acids of hepatitis E virus (HEV) Mexico strain	
CC	CC	capsid protein (AAR96090). It is obtd. by PCR amplification of HEV Mexico	
CC	CC	ORF-2 (AAT27108). The amplified fragment is inserted into vector pGEX for	
CC	CC	prodn. 62K in E. coli cells. Expression of full-length ORF-2 in Sf9	
CC	CC	insect cells using a baculovirus system also results in prodn. of 62K	
CC	CC	antigen (see also AAR96101 and AAR96103). 62K represents an improved	
CC	CC	antigen, in comparison to bacterial expressed proteins, for use in HEV	
CC	CC	diagnostic assays, and also has excellent immunogenic properties.	
CC	CC	(Updated on 16-OCT-2003 to standardise OS field)	
XX	SQ	Sequence 1647 BP; 327 A; 497 C; 400 G; 423 T; 0 U; 0 Other;	
Query March		81.9%; Score 1621.4; DB 2; Length 1647;	
Best Local Similarity		99.5%; Pred. No. 0;	
Matches 1638; Conservative		0; Mismatches 6; Indels 3; Gaps 1;	
Qy	334	GCTGTGGCGCCTGCCCATGACACCTCACCCGTCGGAGCTTGATTCTCGCGGTGCAATT	393
Db	1	GCTGTGGCGCCTGCCCATGACACCTCACCCGTCGGAGCTTGATTCTCGCGGTGCAATT	60
Qy	394	CTACGCCGCCAGTATTAATTGTCTACTTCACCCCTGACATCCTCTGTGGCCTCTGGCACT	453
Db	61	CTACGCCGCCAGTATTAATTGTCTACTTCACCCCTGACATCCTCTGTGGCCTCTGGCACT	120

QY	454	AATTTAGTCTGTATGACAGCCCCCTTAATCGCCTCTCGCGCTGCAGACGGTACTAAT	513	QY	1531	CGATCGCTTGACTGTGTCCTCAAAAGTCACCCCTCGACGGCGGCCCTCCCGACTGTTGAGCAA	1590
Db	121	AATTTAGTCTGTATGACAGCCCCCTTAATCGCCTCTCGCGCTGCAGACGGTACTAAT	180	Db	1201	CGATCGCTTGACTGTGTCCTCAAAAGTCACCCCTCGACGGCGGCCCTCCCGACTGTTGAGCAA	1260
QY	514	ACTCACATTATGCCACACAGAGGCTCCAAATTATGACAGTACCGGTTGCGCGCTACT	573	QY	1591	TATTCACAAGACATCTTTTGTGCTCCGCCCTTCGTGGCAAGCTCTCTTTTGGAGGCGCGC	1650
Db	181	ACTCACATTATGCCACACAGAGGCTCCAAATTATGACAGTACCGGTTGCGCGCTACT	240	Db	1261	TATTCACAAGACATCTTTTGTGCTCCGCCCTTCGTGGCAAGCTCTCTTTTGGAGGCGCGC	1320
QY	574	ATCCGTTACCGGCCCTAGTGCCCTAATGCAAGTTGGAGGCTATGCTATATCCATTTCTTTC	633	QY	1651	ACACAACAAAGCAGGTTATCTTATTAATTAATACTACTGCTAGTGACCAAGATTCTGATT	1710
Db	241	ATCCGTTACCGGCCCTAGTGCCCTAATGCAAGTTGGAGGCTATGCTATATCCATTTCTTTC	300	Db	1321	ACAACAAAAGCAGGTTATCTTATTAATTAATACTACTGCTAGTGACCAAGATTCTGATT	1380
QY	634	TGCGCTCAACACACACACCCCTACATCTGTTGACATGAATTCATTTCTCCACTGAT	693	QY	1711	GAAATGCTGCGCGCCATCGGGTCGCCATTTCAACCTATACACACAGGCTTGGGCGCGGT	1770
Db	301	TGCGCTCAACACACACACCCCTACATCTGTTGACATGAATTCATTTCTCCACTGAT	360	Db	1381	GAAATGCTGCGCGCCATCGGGTCGCCATTTCAACCTATACACACAGGCTTGGGCGCGGT	1440
QY	694	GTCAAGATTCTTGTTCACACTGGCATAGCATCTGAATTTGGTTCATCCCAAGCAGCGCCTT	753	QY	1771	CGGTCGCCATTTCTCGGCGCGGTTTGGTTCACGCTCCGCTCCGCTCTGCTGCTGGAG	1830
Db	361	GTCAAGATTCTTGTTCACACTGGCATAGCATCTGAATTTGGTTCATCCCAAGCAGCGCCTT	420	Db	1441	CGGTCGCCATTTCTCGGCGCGGTTTGGTTCACGCTCCGCTCCGCTCTGCTGCTGGAG	1500
QY	754	CACTACCGCAATCAAGGTTGGCGCTCGGTTGAGACATCTGGTGTGCTGAGGAGGAGCC	813	QY	1831	GATACCTTTTGTATTCGCGGCGGCGCACACATTTGATGACTTCTGCCCTGAATGCCGC	1890
Db	421	CACTACCGCAATCAAGGTTGGCGCTCGGTTGAGACATCTGGTGTGCTGAGGAGGAGCC	480	Db	1501	GATACCTTTTGTATTCGCGGCGGCGCACACATTTGATGACTTCTGCCCTGAATGCCGC	1560
QY	814	ACCTCGGTCCTGTCAATGTTATGATACATACATGGCTCTCCAGTTAACTCCTATACCAATACC	873	QY	1891	GCTTTAGGCTCCAGGTTGTGCTTTCCAGTCAACTGTGCTGAGCTCCAGCGCCTTAAA	1950
Db	481	ACCTCGGTCCTGTCAATGTTATGATACATACATGGCTCTCCAGTTAACTCCTATACCAATACC	540	Db	1561	GCTTTAGGCTCCAGGTTGTGCTTTCCAGTCAACTGTGCTGAGCTCCAGCGCCTTAAA	1620
QY	874	CCTTATACCGGTCCTTGGCTTACTGGACTTTGGCTTTAGAGCTTGAGTTTCCCAATCTC	933	QY	1951	GTTAAGTGGTAAACTCGGAGTTG 1977	
Db	541	CCTTATACCGGTCCTTGGCTTACTGGACTTTGGCTTTAGAGCTTGAGTTTCCCAATCTC	600	Db	1621	GTTAAGTGGTAAACTCGGAGTTG 1647	
QY	934	ACCACCTGTAAACCAATACAGTGTGTCCGTTACTCCAGCACTGCTGCTAC---TCC	990	RESULT 14			
Db	601	ACCACCTGTAAACCAATACAGTGTGTCCGTTACTCCAGCACTGCTGCTAC---TCC	660	AAV61688			
QY	991	GCCGAGGCGCCACCGGACTCGGAGCTGCACCAACTGCAGCCACAGGTTTCATGAAA	1050	ID AAV61688 standard; DNA; 1983 BP.			
Db	661	GCCGAGGCGCCACCGGACTCGGAGCTGCACCAACTGCAGCCACAGGTTTCATGAAA	720	XX AAV61688;			
QY	1051	GATCTCCACTTTACCGGCTTAATGGGTTAGGTGAAGTGGCGCGGATAGCTCTAACA	1110	AC AC			
Db	721	GATCTCCACTTTACCGGCTTAATGGGTTAGGTGAAGTGGCGCGGATAGCTCTAACA	780	DT 03-DEC-1998 (first entry)			
QY	1111	TTACTTAACTTCTGACAGCTCCTCGCGGCGCTCCCGACAGAAATTAATTTCTGCGCT	1170	DE Hepatitis E virus hollow particle protein DNA #2.			
Db	781	TTACTTAACTTCTGACAGCTCCTCGCGGCGCTCCCGACAGAAATTAATTTCTGCGCT	840	XX Hollow particle protein; virus; antibody; detection; immunoassay;			
QY	1171	GCGGGCAACTGTTTATTCGCGCGGTTGTCTCAGCCCAATGCGAGCCAAACCGTGAAG	1230	KW infection; ss.			
Db	841	GCGGGCAACTGTTTATTCGCGCGGTTGTCTCAGCCCAATGCGAGCCAAACCGTGAAG	900	XX Hepatitis virus.			
QY	1231	CTCTATACATCAGTGGAGATGCTCAGCAGGATAAGGGTTGCTATFCCCCACGATATC	1290	XX JP10234383-A.			
Db	901	CTCTATACATCAGTGGAGATGCTCAGCAGGATAAGGGTTGCTATFCCCCACGATATC	960	XX 08-SBP-1998.			
QY	1291	GATCTTGGTATTCGCGTGGTTCATTCAGGATTTATGACACCAACGATGAGCAGATCGG	1350	XX 28-FEB-1997; 97JP-00062445.			
Db	961	GATCTTGGTATTCGCGTGGTTCATTCAGGATTTATGACACCAACGATGAGCAGATCGG	1020	XX 28-FEB-1997; 97JP-00062445.			
QY	1351	CCACCCCGTCGCTCGGCCATCTCGGCCCTTTTCTGTTCTCCGAGCAAAATGATGACTT	1410	XX (DENK-) DENKA SEIKEN KK.			
Db	1021	CCACCCCGTCGCTCGGCCATCTCGGCCCTTTTCTGTTCTCCGAGCAAAATGATGACTT	1080	PA (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.			
QY	1411	TGGGTGCTCCTACTGACGCGAGTATGACAGTCCACTTACGCGTCTCAACTGGCCCG	1470	XX WPI; 1998-535037/46.			
Db	1081	TGGGTGCTCCTACTGACGCGAGTATGACAGTCCACTTACGCGTCTCAACTGGCCCG	1140	XX P-PSDB; AAW76369.			
QY	1471	GTTTATATCTCGGACAGCGTACTTTGCTGAATGTTGCGACTGCGCGCAGGCGGTAGCC	1530	XX Hepatitis E virus hollow particle poly:peptide(s) and nucleic acids			
Db	1141	GTTTATATCTCGGACAGCGTACTTTGCTGAATGTTGCGACTGCGCGCAGGCGGTAGCC	1200	PT encoding it - useful for more accurate detection of HEV in samples, using			
				PT immuno-assays and nucleic acid hybridisation.			
				XX Claim 13; Page 24-26; 29pp; Japanese.			
				CC This sequence encodes a Hepatitis E viral hollow particle protein. This			
				CC polypeptide can be used to raise antibodies to detect HEV infection in			
				CC samples, e.g. by immuno-assay based techniques, and the nucleic acid can			
				CC be used for the same in nucleic acid hybridisation assays. The			
				CC polypeptides and nucleic acids allow more accurate detection of HEV than			

Db 6157 CTCACACCCAGGCTGCTACTCGCTTCATGAAGGACCTCTATTTTACTAGTACTAAUGGT 6216
Qy 1078 GTAGTGAAGTCCGGCCGGGATAGCTCTAACATTTACTTAACCTTCTGACACGCTCCTC 1137
Db 6217 GTTGGTGAGATCGGCCGGGATAGCGCTTACCTGTTTAACTTGTGACACCCCTGCTT 6276
Qy 1138 GCGGGCTCCCGACAGAAATAATTTGCTGGCTGGCGGCAACTGTTTTTATTTCCGCCCG 1197
Db 6277 GCGGCTTACCGACAGAAATTTGCTGGCTGGCGGCAAGTGTCTACTCTCGCCCC 6336
Qy 1198 GTTGTCTCAGCAATGGCGACCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAG 1257
Db 6337 GTCGTCTCAGCAATGGCGACCGACTGTTAAGCTGTATACATCTGTGGAGAAATGCTCAG 6396
Qy 1258 CAGGATAGGGTGTGCTATATCCCCACAGATTCGATCTTGGTGATTCGCGTGTGGTCAATT 1317
Db 6397 CAGGATAGGGGTAATTGCAATCCCGCATGACATCGACCTCGGGGAATCCCGTGTAGTTATT 6456
Qy 1318 CAGGATTATGACAAACAGCATGAGCAGGATCGGCCCAACCCCGTGGCCCTTCCTCGG 1377
Db 6457 CAGGATTATGACAAACCAACATGAGCAGGACCGACCGACACCTTCCCGACGCCCATCGCGT 6516
Qy 1378 CCTTTTCTGTTCTCCGAGCAAAATGATGACTTTTGGTGTCCCTCACTGCAGCCGAGTAT 1437
Db 6517 CCTTTTCTGTTCTCCGAGCTAAGCATGTGCTTTGGCTTCTCTCACCGCTGCCGAGTAT 6576
Qy 1438 GACCAGTCCACTTACGGGTCTCACTGGCCGGTTTATATCTCGGACAGCGTGACTTTG 1497
Db 6577 GACCAGTCCACTTACGGGTCTTTCGACCGGCCCGAGTCTATGCTCTGACTCTGTGACCTTG 6636
Qy 1498 GTGAATGTTGCGACTGGCGCGCAGCCCGTAGCCGATCGCTTGACTGGTCCAAAAGTCACC 1557
Db 6637 GTTAATGTTGCGACCGCGCGCAGGCCGTTCGCCGGTCACTCGACTGACCAAGTCACA 6696
Qy 1558 CTCGACGGCGGCCCTCCCGACTGTTGAGCAATATCCAAAGACATTTTGTGTCGCC 1617
Db 6697 CTTGATGTCGCCGCCCTTTCACCATCCAGCAGTATTCAAAAGACCTTCTTTGTCTGCCG 6756
Qy 1618 CTTGTCGCGAGCTCTCCTTTTGGAGCCCGGCACACAAAAGCAGGTTATCCTTTATAAT 1677
Db 6757 CTCGCGGTAAGCTCTCCTTTTGGAGCGCAGGAACACTAAAGCGGGGTACCCCTTATAAT 6816
Qy 1678 TATAATACTACTGTAGTGACAGATTCTGATTGAAAATGTCGCCGCCATCGGCTGCC 1737
Db 6817 TATAACACCACTGTAGTGACCAACTGCTCGTTGAGATGCGCTGGGCATCGGTTGCT 6876
Qy 1738 ATTTCAACCTATACACAGGCTTGGGCGCGGTCCGFTCGCATTTTCTGCGGCCGCCGTT 1797
Db 6877 ATTTCCACCTACACTACTAGCTGGGTGCTGGCCCGGTCTCTAATTTCCGCGGTTGCTGTT 6936
Qy 1798 TTGGCTCAGCTCGGCCCTGGCTCTGCTGGAGGATCTTTTGATTATCCGGGGCGGGCG 1857
Db 6937 TTAGCCCCCACTCTGTGTAGCATTTGCTTGAGGATACCATGGAATACCCCTGCCCGGCC 6996
Qy 1858 CACACATTTGATGACTTCTGCCCTGAATGCCGCGCTTTAGGCCCTCCAGGGTTGTCTTTC 1917
Db 6997 CATACTTTCGATGACTTCTGCCCGAGTGCGCCGCCCTTGGCCCTCCAGGGTTGTCTTTT 7056
Qy 1918 CAGTCAACTGTCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAAACCTCGGAGTTG 1977
Db 7057 CAGTCTACTGTCTGAGCTTTCAGCGCCTTAAGATGAAGTGGGTAAAACCTCGGAGTTA 7116
Qy 1978 TAG 1980
Db 7117 TAG 7119

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 10:14:02 ; Search time 5109.13 Seconds
(without alignments)
11572.836 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 1980

Sequence: 1 ATGGCCCTAGGCTCTTTT.....GTAAACTCGGAGTTGTAG 1980

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1.*
10: gb_est2.*
11: gb_hc:*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_fum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vit.*
21: em_gss_fun.*
22: em_gss_nam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62.4	3.2	925	29	CNS0091P
C 2	61.8	3.1	1201	13	BX360624
C 3	60.2	3.0	844	29	CNS0052P
C 4	59.4	3.0	1103	13	BX403654

5	58.2	2.9	1013	29	CNS016KT
C 6	57.6	2.9	1000	13	BX407619
C 7	57.4	2.9	1101	29	CNS017SY
C 8	57.2	2.9	1203	29	CNS015Y4
9	56.6	2.9	828	13	BX425797
10	54.8	2.8	1201	13	BX381961
C 11	54	2.7	1201	9	AL574009
C 12	53.8	2.7	839	29	CNS004NB
C 13	53.8	2.7	1193	13	BX421125
C 14	53.2	2.7	993	13	BQ144482
15	53	2.7	1201	13	BX356664
C 16	52.6	2.7	1144	13	BX415926
17	52.4	2.6	807	13	BX464554
C 18	52	2.6	1201	9	AL538546
C 19	51.8	2.6	1057	13	BX349688
C 20	51.8	2.6	1201	13	BX335523
C 21	51.2	2.6	1036	9	AL550172
C 22	50.8	2.6	1201	29	CNS015WQ
C 23	50.6	2.6	877	12	B1956655
C 24	49.2	2.5	1009	29	CNS010EW
C 25	49	2.5	975	23	CNS02KED
C 26	49	2.5	1100	29	CNS001BC
C 27	49	2.5	1201	9	AL525224
C 28	48.8	2.5	935	29	CNS006XK
C 29	48.6	2.5	1189	13	BX407485
C 30	48.4	2.4	863	12	B1953455
C 31	48.2	2.4	882	28	BZ050854
C 32	48.2	2.4	1061	13	BX391246
33	48.2	2.4	1071	29	CNS000EM
34	48.2	2.4	1101	29	CNS00LXJ
C 35	48	2.4	359	14	CB908167
C 36	47.8	2.4	995	29	AG060386
C 37	47.8	2.4	1101	29	CNS016HG
C 38	47.8	2.4	1160	29	AG043473
C 39	47.8	2.4	1201	13	BX340485
C 40	47.6	2.4	689	29	AG044460
C 41	47.6	2.4	906	29	AG081338
C 42	47.6	2.4	1136	28	CC292742
C 43	47.4	2.4	1184	13	BX425067
C 44	47.2	2.4	989	29	AG063827
C 45	47.2	2.4	1273	12	BM562099

ALIGNMENTS

RESULT 1	CNS0091P	925 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
DEFINITION	BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL053013				
VERSION	AL053013.1				
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 925)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosier in Pieter de Jong's laboratory in the Department of				

double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

```

FEATURES
source
Location/Qualifiers
1..844
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR11p15"
/clone_lib="RPEC198"
/note="end . 7571"

```

```

ORIGIN
Query Match 3.0%; Score 60.2; DB 29; Length 844;
Best Local Similarity 19.4%; Pred. No. 0.0018;
Matches 59; Conservative 131; Mismatches 114; Indels 0; Gaps 0

23 TGTGTTCCTCTGTGTTTCCTATGTTGCGCGGCCACGACGGTCAGCGGTCGTGCC 82
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
676 TCWTHTNSNTYTWVYVYVTHHTYVNTCCYYCCGCCCYCSCBYBCNTCSYYSVBS 617
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
83 GCGCTGCTGGCGCGCAGCGCGGTACCGCGCGGTGTTTCTGGGGTACCGGGTTGATT 142
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
616 GCSSSSSSCCSSSSSHYSCCBBCCBGCGCYBVCSSBSBTBSSVGSBSSTSGTGCK 557
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
143 CTCAGCCCTTCGAATCCCTATATTCATCCAAACCAACCCCTTTGCCCCACAGCTGCG 202
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
556 CBSSSSSSBSCSSBSSCCCYCSSYCTCYVSSCCBYCCBSGCBTBSVBSCSSGHTBSSBSG 497
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
203 CTCGGTCGGGTCTGGACCTCGCTTCGCCAAACAGCCCGCGGCACTTGCGTCCACTGGC 262
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
496 GKSGGKSSRGGBGKSGGKYCYCBTYCCBSBSBCCSSSSSTSKCCSGSSSTKCGSSB 437
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
263 GAGATCAGGCCCGAGCGCCCTCCGCTGCCTCCGCTGCGCCACCTGCCACACCGGGGTG 322
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
436 BGGGSSSKTSGSSSCBSSSSSTSGSSGSKKYSYSTSSSSCGGTGKNTTTCGVGSTCTTT 377
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
323 CGGC 326
   ||: ||
376 KSSC 373
   ||: ||
Db

```

RESULT 4	1103 bp	mRNA	linear	EST 15-MAY-2003
LOCUS	1103 bp	mRNA	linear	EST 15-MAY-2003
DEFINITION	1103 bp	mRNA	linear	EST 15-MAY-2003
ACCESSION	1103 bp	mRNA	linear	EST 15-MAY-2003
VERSION	1103 bp	mRNA	linear	EST 15-MAY-2003
KEYWORDS	1103 bp	mRNA	linear	EST 15-MAY-2003
SOURCE	1103 bp	mRNA	linear	EST 15-MAY-2003
ORGANISM	1103 bp	mRNA	linear	EST 15-MAY-2003
REFERENCE	1103 bp	mRNA	linear	EST 15-MAY-2003
AUTHORS	1103 bp	mRNA	linear	EST 15-MAY-2003
TITLE	1103 bp	mRNA	linear	EST 15-MAY-2003

Genoscope - Centre National de Sequençage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA007ZH12Pp1.
Location/Qualifiers
1 1103

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Cl0BA007ZH12"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched.

```

ORIGIN

Query Ma
Best Loc
Matches

QY Db

QY Db

Db

DB
ON

Db Qy

Db

CNS016KT
LOCUS
DEFINITION

ACCESSION
VERSION
KEYWORDS

ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
SOURCE

ORIGIN

double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

3.0%; Score 59.4; DB 13; Length 1103;
Identity 28.0%; Pred. No. 0.0033;
Conservative 93; Mismatches 151; Indels 0; Gaps 0;

TCCTGTTTCTGCTATGTGCCCGCCACCGACCGTGAGCCGTCTGGCCGC 84
||||| : ||||| : : : : : : : : : : : : : : : : : : :
||||| : ||||| : : : : : : : : : : : : : : : : : : :
TTTTTTTTTDRGRPWTCTCSSSSSTSSSGSSSSSSSBSGSSSSSS 467

SSGGCGGCGAGCCGCCGTACCGCGGTGGTTCTTGCGGTACCGGGTGAATTCT 144
SSSSSSSCS SSSSSSSSSSSSSSSSSSCSSSGSSSSSSSSSSSSSSSG 407

: ||: | || | ||: ||| | : : : :
 SSSGGGCCCCGCGCCCCCCCCCCCCCCCCCCCCSSSSSSSSSG 347

[illegible][illegible]

CCCTYYTCCSCCCSCCCMCCCCCCCCYYMMNNC 188

1013 bp DNA linear GSS 26-JUL-1999
a melanogaster genome survey sequence SP6 end of BAC
of DrosBAC library from Drosophila melanogaster (fruit

1 GI:5624218

a melanogaster
; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Endopterygota; Diptera; Brachycera; Muscomorpha;

Submission
1 (23-JUL-1999) Genoscope - Centre National de Séquençage

tion of this BAC-end sequence was carried out as part of a
tion with the European Drosophila Genome Project (EDGP) -

Dros BAC) was made by Alain Billaud at CEPH (Centre National de la Recherche Scientifique, France) and the *lacZ* Polymorphisme Humain) with funding provided by a MRC grant. The DNA was prepared from embryos by Alain Bucheton (University of Cambridge, UK). The *lacZ* reporter gene construct has been described previously (Billaud et al., 1990).

```
1.
Location/Qualifiers
1. .1013
/organism="Drosophila melanogaster"
```

```

..._lib="genomic bank"
/db_xref="taxon:7227"
/clone="BACN16J16"
/clone_lib="DrosBAC"

```

```
/note="end : SP6"
```


TITLE Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:12933798.
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10873.f
For more information about this cluster, see
<http://www.genoscope.cns.fr/>
cgi-bin/Cluster.cgi?seq=CS0DI053BA02NP1&cluster=10873.f. Contact :
Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DI053BA02NP1.

```

FEATURES
source
railway Avenue Genoscope sequence ID : CSOD1053BAU2NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1053YA04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

Query Match	2.7%;	Score 54;	DB 9;	Length 1201;
Best Local Similarity	39.2%;	Pred. No. 0.07;		
Matches	94;	Conservative 40;	Mismatches 106;	Indels 0; Gaps 0
QY	91	GGCGCGCGCAGCGCGGTACCGCGCGTGTTCCTGGCGTGACCGGGTGTATCTTCAGCC	150	
DB	957	TTTGSGTGTGTGCGGGCGGGGCGGGTWTATCCMCMGGGSATTGCCGKTCKMSGGCC	898	
QY	151	TTCGCAATCCCTATATTTCACCAACACCCCTTTGCCCCACAGCTTGCCGTGCGTCC	210	
DB	897	GGGSGGGCCSCMGTTTGGCGGCCCMSCMGMGTTGTTGCAGCMGWTGMCAGTGGCA	838	
QY	211	GGGTCTGGACCTCGCTTTTCGCCAACACCGCCGGCCACTTGGCTCCACTTGGCGAGATCA	270	
DB	837	SMGGCATACMCGMMTTTCMGGSGCGTGTGCMGCTCTGGGCGCGCTCGMGCMCTGC	778	
QY	271	GCCACAGCGCCCTCGCTGCTCCCGTGGCGACCTGCCACACCGCGGGCTCGCGCGTG	330	
DB	777	GSCMGSGCCCTGTARGCGKTTSCCMGGSGCCCTTGSSCSGSGSCCTGSCCTHTCS	718	

RESULT 12	CNS004NB/c	839 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	CNS004NB/c				
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BAC10816 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL054280				
VERSION	AL054280.1	GI:4931788			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1 (bases 1 to 839)				
TITLE	Genoscope.				
JOURNAL	Direct Submission				
COMMENT	Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila				

melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial, EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila/bac.htm>.

FEATURES
source

ORIGIN

Query Match	2.7%;	Score 53.8;	DB 29;	Length 839;
Best Local Similarity	18.6%;	Pred. No. 0.067;		

Qy	2	TGGCGCCTAGGCGCTCTTTTGTCTGTGTCTCTCTTTCTGCCTATGTTCGCCGCGCCAC	61
Db	819	TSSSSTSTSTSKTGCKSKGSBSGSSSSSSGCGSTYSSSTGGSGSVSVSGSSGSSBGS	760
Qy	62	CGACCGGTACGCCGTCTGGCGCCGTGTGGCGGCGCAGGCGCGGTACCGCGGTGGTT	121
Db	759	SGSGBSGCSGSSSGSSGSKSSSGSGSGSCSSSGCGGCGSTCSSTTCSSTSGSKSSG	700
Qy	122	TCTGGGGTGACCGGGTGTGATTTCTCAGCCCTTCGCAATCCCTATATCATCAACCAAC	181
Db	699	CSGCSSTSTGBSGTGYTTCSTSSCTSCSSGTTTCSBSTTSSSTTSTTTSBTTC	640
Qy	182	CTTTTGGCCACAGCGTTGCGCTCGGTCTGGACCTCGCCTTGCGCAACACGACC	241
Db	639	STTTCKCKCYTCVTBT	580
Qy	242	GGCCACTTGGCTTCCACTTGGCGAGATCAGGCCACAGCCCTCGCTGCTCCCGTCGCC	301
Db	579	CTSSSSSTSTSSGSHSSSSTCBTCGCGSCYSCCCSCSCTCCSCSYSCSTCCSCC	520
Qy	302	GACCTGCCACAGCCGGGCGCTCGCGGTGACGGCTGTGGCCCGCTG	346
Db	519	TSSSSCTCGSCSTSSSCTTCTCSTGTSSGTTBGSSTSGSSS	475

RESULT 13	ACCESSION	REFERENCE
EX421125/c	VERSION	AUTHORS
LOCUS	KEYWORDS	TITLE
DEFINITION	SOURCE	JOURNAL
	ORGANISM	COMMENT

BX421125 1193 bp mRNA linear EST 15-MAY-2003
 BX421125 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
 clone CS02G001YM04 3-PRIME, mRNA sequence.
 BX421125
 BX421125.1 GI:30768081
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1193)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1904.f For

more information about this cluster, see
[http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DG001BG02NP1&cluster=3904.f](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DG001BG02NP1&cluster=3904.f). Contact :
 Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com> / Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DG001BG02NP1

FEATURES source

1. l1193
 location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="CS0DG001Y004"
 /tissue_type="B CELLS (R
 /cell_lines="RAMOS CELL L
 /clone_libs="Homo sapiens
 /note="Vector: pCMVSPORT
 with a NotI-oligo(dT) pr
 double-strand cDNA was
 the Not I and EcoRV site
 Library was not normalized

ORIGIN

	Query Match	2.7%; Score 53.8; DB 13; Length 1193;
	Best Local Similarity	37.2%; P-red. No. 0.079;
	Matches	Conservative 41; Mismatches 182; Indels 0; Gaps 0;
a	aa	
b	bb	
c	cc	
d	dd	
e	ee	
f	ff	
g	gg	
h	hh	
i	ii	
j	jj	
k	kk	
l	ll	
m	mm	
n	nn	
o	oo	
p	pp	
q	qq	
r	rr	
s	ss	
t	tt	
u	uu	
v	vv	
w	ww	
x	xx	
y	yy	
z	zz	

RESULT 14

Accession	BQ144482	993 bp	mRNA	linear	EST 24-APR-2002
LOCUS	BQ144482				
DEFINITION	NF085H1DPT1F1095 Drought Medicago truncatula cDNA clone NF085H1DPT1F1095, mRNA sequence.				
ACCESSION	BQ144482				
VERSION	BQ144482.1 GI:20281541				
KEYWORDS	EST.				
SOURCE	Medicago truncatula (barrel medic)				
ORGANISM	Medicago truncatula				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.				
REFERENCE	1. (bases 1 to 993)				
AUTHORS	Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.				
TITLE	Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula drought library				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: May GD Plant Biology Division				

The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmaw@noble.org
Insert Length: 993 Std Error: 0.00
plate: 085 row: H column: 11
Seq primer: TCACACGAGAACAGCTATGAC.

FEATURES	SOURCE
1. Feature 1	Source 1
2. Feature 2	Source 2
3. Feature 3	Source 3
4. Feature 4	Source 4
5. Feature 5	Source 5
6. Feature 6	Source 6
7. Feature 7	Source 7
8. Feature 8	Source 8
9. Feature 9	Source 9
10. Feature 10	Source 10
11. Feature 11	Source 11
12. Feature 12	Source 12
13. Feature 13	Source 13
14. Feature 14	Source 14
15. Feature 15	Source 15
16. Feature 16	Source 16
17. Feature 17	Source 17
18. Feature 18	Source 18
19. Feature 19	Source 19
20. Feature 20	Source 20
21. Feature 21	Source 21
22. Feature 22	Source 22
23. Feature 23	Source 23
24. Feature 24	Source 24
25. Feature 25	Source 25
26. Feature 26	Source 26
27. Feature 27	Source 27
28. Feature 28	Source 28
29. Feature 29	Source 29
30. Feature 30	Source 30
31. Feature 31	Source 31
32. Feature 32	Source 32
33. Feature 33	Source 33
34. Feature 34	Source 34
35. Feature 35	Source 35
36. Feature 36	Source 36
37. Feature 37	Source 37
38. Feature 38	Source 38
39. Feature 39	Source 39
40. Feature 40	Source 40
41. Feature 41	Source 41
42. Feature 42	Source 42
43. Feature 43	Source 43
44. Feature 44	Source 44
45. Feature 45	Source 45
46. Feature 46	Source 46
47. Feature 47	Source 47
48. Feature 48	Source 48
49. Feature 49	Source 49
50. Feature 50	Source 50
51. Feature 51	Source 51
52. Feature 52	Source 52
53. Feature 53	Source 53
54. Feature 54	Source 54
55. Feature 55	Source 55
56. Feature 56	Source 56
57. Feature 57	Source 57
58. Feature 58	Source 58
59. Feature 59	Source 59
60. Feature 60	Source 60
61. Feature 61	Source 61
62. Feature 62	Source 62
63. Feature 63	Source 63
64. Feature 64	Source 64
65. Feature 65	Source 65
66. Feature 66	Source 66
67. Feature 67	Source 67
68. Feature 68	Source 68
69. Feature 69	Source 69
70. Feature 70	Source 70
71. Feature 71	Source 71
72. Feature 72	Source 72
73. Feature 73	Source 73
74. Feature 74	Source 74
75. Feature 75	Source 75
76. Feature 76	Source 76
77. Feature 77	Source 77
78. Feature 78	Source 78
79. Feature 79	Source 79
80. Feature 80	Source 80
81. Feature 81	Source 81
82. Feature 82	Source 82
83. Feature 83	Source 83
84. Feature 84	Source 84
85. Feature 85	Source 85
86. Feature 86	Source 86
87. Feature 87	Source 87
88. Feature 88	Source 88
89. Feature 89	Source 89
90. Feature 90	Source 90
91. Feature 91	Source 91
92. Feature 92	Source 92
93. Feature 93	Source 93
94. Feature 94	Source 94
95. Feature 95	Source 95
96. Feature 96	Source 96
97. Feature 97	Source 97
98. Feature 98	Source 98
99. Feature 99	Source 99
100. Feature 100	Source 100

```

1. 993
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF085H11PT"
/tissue_type="Plantlets"
/dev_stage="Pooled timepoints"
/clone_lib="Drought"
/note="Vector: Lambda Zap; Cont
plantlets harvested in a series
timepoints."

```

ORIGIN

Query Match	2.78;	Score 53.2;	DB 13;	Length 993;
Best Local Similarity	48.7%;	Pred. No. 0.17		
Matches 145;	Conservative 0;	Mismatches 153;	Indels 0;	Gaps 0;
Qy	75	GTCTGGCGCGCGTCTGGCGGGCGACAGCGCGGTACCGCGGGTGGTTCTTGGGGTGACCG	134	
Db	79	GGCGGAGGCGCGCGGGGGCGCGCGGGGAGGGCGCGGGGGCGCGGGCGGGCGG	138	
Qy	135	GGTTGATCTCAGCCCTTTCGCAATCCCTATATTCATCCAAACACCCCTTTGCCCCAGA	194	
Db	139	CGCGGCCCCGACAGCGGCGCCCGCGCCCCCAAGATCGCCCGCGGCCACCTTTTCCCCCCC	198	
Qy	195	CGTTGCGCGTGCGTGCGGGTCTGGAACCTCGCTTTCGCGCAACAGCGCGGCGACATTGGGTC	254	
Db	199	CCGTCCCTTCTATCTCGGCGCCCCCTCCGCCCTCGCTCGCTCCGCCCTCTTTCGCGCCC	258	
Qy	255	CACTTGGGAGATCAGGCGCCAGCGCCCTCGGTGCGCTCCGTCGCGCACTGCCACAGC	314	
Db	259	CCTCCCCCTCCCCCCCCCCCCCCCCCTTAGCCTCTCTCTGGGCCCGGCTTCGCCCCC	318	
Qy	315	CGGGGCTGCGGCGCTGACGGCTGTGGCGCTTGCCTATGACACCTGACCCCGTCCCGGAC	372	
Db	319	CCGGCCCCCGGCCCCCGGCCCCCCCCCCCCCGCGCCGCGCCCCCGCGCGCCCGCTC	376	

RESULT 15
BX356664

EX356664	1201 bp	mrna	linear	EST 05-MAY-2000
DEFINITION	EX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSODI015YB03 3-PRIME, mRNA sequence.			
ACCESSION	EX356664			
VERSION	EX356664.1	GI:30378083		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequençage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr , Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS00D1015CA02NP1. location/Qualifiers			
FEATURES	1. .1201			
source				

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 10:24:37 ; Search time 140.394 Seconds
(without alignments)
7826.585 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 1980

Sequence: 1 ATGCGCCCTAGGCTCTTTT.....GTAAACTCGGAGTGTAG 1980

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1980	100.0	1980	1	US-08-240-049B-2
2	1980	100.0	2055	1	US-08-259-148A-4
3	1980	100.0	2055	1	US-08-484-054-4
4	1980	100.0	2055	2	US-07-876-941A-4
5	1980	100.0	2055	3	US-08-477-292-2
6	1980	100.0	2055	4	US-07-870-985A-4
7	1980	100.0	2100	1	US-08-259-148A-2
8	1980	100.0	2100	1	US-08-484-054-2
9	1980	100.0	2100	2	US-07-876-941A-2
10	1980	100.0	2100	4	US-07-870-985A-2
11	1980	100.0	7171	3	US-08-478-507-10
12	1980	100.0	7171	3	US-09-128-275A-10
13	1980	100.0	7171	4	US-09-553-427-10
14	1957.4	98.9	2058	3	US-08-542-634-2
15	1957.4	98.9	2058	5	PCT-US95-13703-2
16	1644	83.0	1644	3	US-08-477-292-4
17	1621.4	81.9	1647	3	US-08-542-634-4
18	1621.4	81.9	1647	5	PCT-US95-13703-4
19	1378.2	69.6	7168	3	US-08-840-316-4
20	1378.2	69.6	7168	4	US-08-809-523-4
21	1378.2	69.6	7168	3	US-08-471-971-4
22	1378.2	69.6	7168	4	US-09-402-776-4
23	1378.2	69.6	7168	5	PCT-US93-08849A-4
24	1378.2	69.6	7168	5	PCT-US93-08849A-4
25	1371.8	69.3	1984	4	US-09-172-699-1
26	1368.6	69.1	1983	1	US-08-240-049B-1
27	1368.6	69.1	2049	1	US-08-259-148A-3

28 1368.6 69.1 2049 1 US-08-484-054-3 Sequence 3, Appli
29 1368.6 69.1 2049 2 US-07-876-941A-3 Sequence 3, Appli
30 1368.6 69.1 2049 3 US-08-542-634-1 Sequence 1, Appli
31 1368.6 69.1 2049 3 US-08-477-292-1 Sequence 1, Appli
32 1368.6 69.1 2049 4 US-07-870-985A-3 Sequence 3, Appli
33 1368.6 69.1 2049 5 PCT-US95-13703-1 Sequence 1, Appli
34 1368.6 69.1 2094 1 US-08-259-148A-1 Sequence 1, Appli
35 1368.6 69.1 2094 1 US-08-484-054-1 Sequence 1, Appli
36 1368.6 69.1 2094 2 US-07-876-941A-1 Sequence 1, Appli
37 1368.6 69.1 2094 3 US-07-870-985A-1 Sequence 1, Appli
38 1368.6 69.1 7195 3 US-08-478-507-6 Sequence 6, Appli
39 1368.6 69.1 7195 4 US-09-128-275A-6 Sequence 6, Appli
40 1368.6 69.1 7195 4 US-09-553-427-6 Sequence 6, Appli
41 1308 66.1 1308 1 US-08-259-148A-10 Sequence 10, Appli
42 1308 66.1 1308 1 US-08-484-054-10 Sequence 10, Appli
43 1308 66.1 1308 2 US-07-876-941A-10 Sequence 10, Appli
44 1308 66.1 1308 4 US-07-870-985A-10 Sequence 10, Appli
45 1289.2 65.1 7207 4 US-09-462-606-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-240-049B-2
; Sequence 2, Application US/08240049B
; Patent No. 5686239
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Tam, Albert W.
; APPLICANT: Varbough, Patrice O.
; TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,049B
; FILING DATE: 09-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles K. Sholtz
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)
; INDIVIDUAL ISOLATE: ORF-2 region
US-08-240-049B-2

Query Match 100.0%; Score 1980; DB 1; Length 1980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTTAGGCTCTTTTGTGTTGTTCTCTTCTGCTATGTTGCGCGGCA 60
Db |||||
QY 1 ATGGCCCTTAGGCTCTTTTGTGTTGTTCTCTTCTGCTATGTTGCGCGGCA 60
Db |||||
QY 61 CCGACCGGTGACCGGTGCTGCGCGCGCTGCTGCGCGCGCGAGCGGCGGTGCT 120
Db |||||
QY 61 CCGACCGGTGACCGGTGCTGCGCGCGCTGCTGCGCGCGCGAGCGGCGGTGCT 120
Db |||||
QY 121 TTCTGGGTGACCGGTGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCT 180
Db |||||
QY 121 TTCTGGGTGACCGGTGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCT 180
Db |||||
QY 181 CCGTTGCGCGCGAGCGGTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCT 240
Db |||||
QY 181 CCGTTGCGCGCGAGCGGTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCT 240
Db |||||
QY 241 CCGCCACTTGGCTTCCACTTGGCGAGATCAGGCGCGCGCTGCTGCGCGCGCT 300
Db |||||
QY 241 CCGCCACTTGGCTTCCACTTGGCGAGATCAGGCGCGCGCTGCTGCGCGCGCT 300
Db |||||
QY 301 CGACCTGCGACAGCGCGGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCT 360
Db |||||
QY 301 CGACCTGCGACAGCGCGGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCT 360
Db |||||
QY 361 CCGCTGCGCGAGCGGTGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCT 420
Db |||||
QY 361 CCGCTGCGCGAGCGGTGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCT 420
Db |||||
QY 421 TCACCCCTGACACTCTCTGCGCGCTTGGCGCTTAAATTTAGTCTGATGAGCGCGCTT 480
Db |||||
QY 421 TCACCCCTGACACTCTCTGCGCGCTTGGCGCTTAAATTTAGTCTGATGAGCGCGCTT 480
Db |||||
QY 481 AATCGCGCTGCGCGTGCAGGCGGTGCTGCGCGCTGCGCGCGCTGCGCGCGCTTCC 540
Db |||||
QY 481 AATCGCGCTGCGCGTGCAGGCGGTGCTGCGCGCTGCGCGCGCTGCGCGCGCTTCC 540
Db |||||
QY 541 AATTTGCGACGACCGGTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCTTAA 600
Db |||||
QY 541 AATTTGCGACGACCGGTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCTTAA 600
Db |||||
QY 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTCTGCGCTCAAAACCAACCCCTTACA 660
Db |||||
QY 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTCTGCGCTCAAAACCAACCCCTTACA 660
Db |||||
QY 661 TCTGTTGACATGAATTCCTATCTTCACTGATGTCAGGATTTCTTCACTTGGGCTG 720
Db |||||
QY 661 TCTGTTGACATGAATTCCTATCTTCACTGATGTCAGGATTTCTTCACTTGGGCTG 720
Db |||||
QY 721 GCATCTGAATTTGTCATCCAGCGGCGCTTCACTTACCGCAATCAAGGTTGGCGCTG 780
Db |||||
QY 721 GCATCTGAATTTGTCATCCAGCGGCGCTTCACTTACCGCAATCAAGGTTGGCGCTG 780
Db |||||
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGGAGCCACTTCCGCTGCTGTCATGTTATGCA 840
Db |||||
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGGAGCCACTTCCGCTGCTGTCATGTTATGCA 840
Db |||||
QY 841 CATGGCTCTCGAGTTAACTCTATACCAATACCCCTTATACCGGTGCTTGGCTTACTG 900
Db |||||
QY 841 CATGGCTCTCGAGTTAACTCTATACCAATACCCCTTATACCGGTGCTTGGCTTACTG 900
Db |||||
QY 901 GACTTTGCTTAGAGCTTGAATTTGCAATCTCAACCTGTAAACCAATACAGTGTG 960
Db |||||
QY 901 GACTTTGCTTAGAGCTTGAATTTGCAATCTCAACCTGTAAACCAATACAGTGTG 960
Db |||||
QY 961 TCCCGTTACTCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db |||||
QY 961 TCCCGTTACTCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db |||||
QY 1021 ACCCAACTGAGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db |||||
QY 1021 ACCCAACTGAGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db |||||
QY 1081 GGTGAAGTGGCGCGGAGTAGCTCTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1140

Db |||||
QY 1081 GGTGAAGTGGCGCGGAGTAGCTCTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1140
Db |||||
QY 1141 GGGCTCCCGACAGAAATTAATTTGCTGCGGTGCGCGGCAACTGTTTATTTCCCGCGGTT 1200
Db |||||
QY 1141 GGGCTCCCGACAGAAATTAATTTGCTGCGGTGCGCGGCAACTGTTTATTTCCCGCGGTT 1200
Db |||||
QY 1201 GTCTCAGCAATGGCGAGCAACCGTGAAGCTCTATATACATAGTGGAGAACTCAGCAG 1260
Db |||||
QY 1201 GTCTCAGCAATGGCGAGCAACCGTGAAGCTCTATATACATAGTGGAGAACTCAGCAG 1260
Db |||||
QY 1261 GATAAGGTTGTTGCTATCCCCCAAGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 1320
Db |||||
QY 1261 GATAAGGTTGTTGCTATCCCCCAAGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 1320
Db |||||
QY 1321 GATTATGACCAACGAGCATGAGCAGGATCGGCCACCCCGTCCGCTGCGCGCACTCGCGCT 1380
Db |||||
QY 1321 GATTATGACCAACGAGCATGAGCAGGATCGGCCACCCCGTCCGCTGCGCGCACTCGCGCT 1380
Db |||||
QY 1381 TTTTCTGTTCTCCGAGCAAAATGATGTACTTTGGCTGCTCCCTCACTGAGCGGAGTATGAC 1440
Db |||||
QY 1381 TTTTCTGTTCTCCGAGCAAAATGATGTACTTTGGCTGCTCCCTCACTGAGCGGAGTATGAC 1440
Db |||||
QY 1441 CAGTCCACTTACGGGTGCTCAACTGGCCCGGTTTATATCTCGGACAGCGTGAATTTGGTG 1500
Db |||||
QY 1441 CAGTCCACTTACGGGTGCTCAACTGGCCCGGTTTATATCTCGGACAGCGTGAATTTGGTG 1500
Db |||||
QY 1501 AATGTTGCCACTGGCGCGAGCGGTGAGCCCGATCGCTTGAATGCTGCTGCTGCTGCTGCTGCT 1560
Db |||||
QY 1501 AATGTTGCCACTGGCGCGAGCGGTGAGCCCGATCGCTTGAATGCTGCTGCTGCTGCTGCTGCT 1560
Db |||||
QY 1561 GACGGCGCGCTCCCGCTCCGACTGTTGAGCAATATTTCAAGACATTTCTTGTGCTCCCGCTT 1620
Db |||||
QY 1561 GACGGCGCGCTCCCGCTCCGACTGTTGAGCAATATTTCAAGACATTTCTTGTGCTCCCGCTT 1620
Db |||||
QY 1621 CGTGGCAAGCTCTCTTTTGGGAGCGCGGCAAGCAAGAGCAGGTTATCTTATATAT 1680
Db |||||
QY 1621 CGTGGCAAGCTCTCTTTTGGGAGCGCGGCAAGCAAGAGCAGGTTATCTTATATAT 1680
Db |||||
QY 1681 AATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db |||||
QY 1681 AATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db |||||
QY 1741 TCAACCTATACCAAGGCTTGGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db |||||
QY 1741 TCAACCTATACCAAGGCTTGGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db |||||
QY 1801 GCTCCAGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db |||||
QY 1801 GCTCCAGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db |||||
QY 1861 ACATTTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db |||||
QY 1861 ACATTTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db |||||
QY 1921 TCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db |||||
QY 1921 TCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db |||||

RESULT 2
US-08-259-148A-4
; Sequence 4, Application US/08259148A
; Patent No. 5741490
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; APPLICANT: Yarbough, Patrice D.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method

1321 GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCCCGTGCCTCGCCGATCTCGGCT 1380
1321 GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCCCGTGCCTCGCCGATCTCGGCT 1380
1381 TTTTCTGTTCTCCGAGCAATGATGATCTTGGCTGTCCCTCACTGCAGCCGAGTATGAC 1440
1381 TTTTCTGTTCTCCGAGCAATGATGATCTTGGCTGTCCCTCACTGCAGCCGAGTATGAC 1440
1441 CAGTCACTTACGGGTGCTCACTGGCCCGGTTTATATCTCGACAGCGTGATTTGGTG 1500
1441 CAGTCACTTACGGGTGCTCACTGGCCCGGTTTATATCTCGACAGCGTGATTTGGTG 1500
1501 ATGTTGGACGCTGCGCGCAGCGCGTAGCCGATCGCTTGAATGCTGCTCAAAAGTCACCCCTC 1560
1501 ATGTTGGACGCTGCGCGCAGCGCGTAGCCGATCGCTTGAATGCTGCTCAAAAGTCACCCCTC 1560
1561 GACGGCGCGCCCTCCCGACTGTGAGCAATATTCACAGCAATTTCTTGTGTCCCCCTT 1620
1561 GACGGCGCGCCCTCCCGACTGTGAGCAATATTCACAGCAATTTCTTGTGTCCCCCTT 1620
1621 CTTGGCAAGCTCTCTTTTGGGAGCGCGCACAAAGAGCAGTTATCTTATATATAT 1680
1621 CTTGGCAAGCTCTCTTTTGGGAGCGCGCACAAAGAGCAGTTATCTTATATATAT 1680
1681 AATACTACTGCTAGTGACAGATTTCTGATTTGAAATGCTGCGGCCATCGGCTCGCAT 1740
1681 AATACTACTGCTAGTGACAGATTTCTGATTTGAAATGCTGCGGCCATCGGCTCGCAT 1740
1741 TCAACCTATACCAACAGGCTTGGGCGCGTCCGCTCGCATTTCTCGCGCGCGGTTTG 1800
1741 TCAACCTATACCAACAGGCTTGGGCGCGTCCGCTCGCATTTCTCGCGCGCGGTTTG 1800
1801 GCTCCAGCTCCCGCTGCTGCTGAGGATCTTTGATTTATCCGGCGCGCGCAC 1860
1801 GCTCCAGCTCCCGCTGCTGCTGAGGATCTTTGATTTATCCGGCGCGCGCAC 1860
1861 ACATTTGATGACTTCTGCGCTGAATCCGCGCTTTAGGCTCCAGGCTTGTCTTCCAG 1920
1861 ACATTTGATGACTTCTGCGCTGAATCCGCGCTTTAGGCTCCAGGCTTGTCTTCCAG 1920
1921 TCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAGCTCGGGAGTTGTAG 1980
1921 TCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAGCTCGGGAGTTGTAG 1980

RESULT 3

US-08-484-054-4
; Sequence 4, Application US/08484054
; Patent No. 5770689
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krwczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,054
; FILING DATE: 07-JUN-1995

CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; FILING DATE: 20-APRIL-1992
; PRIOR APPLICATION DATA: US 870,985
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Shultz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2055 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 7
; US-08-484-054-4

Query Match 100.0%; Score 1980; DB 1; Length 2055;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCGCCCTAGGCTCTTTGCTGTTGTTGCTCTTCTGCTGCTATGTTGCCGCGCCA 60
DB 1 ATGCGCCCTAGGCTCTTTGCTGTTGTTGCTCTTCTGCTGCTATGTTGCCGCGCCA 60
QY 61 CCGACCGGTGAGCCGCTGTCGCGCGCGGTGTCGCGCGCGGAGCGCGGTACCGCGGTGGT 120
DB 61 CCGACCGGTGAGCCGCTGTCGCGCGCGGTGTCGCGCGCGGAGCGCGGTACCGCGGTGGT 120
QY 121 TTCTGGGGTACCGGGTGTGATTTCTGAGCCCTTCCGCAATCCCTATATTCATCAACCAAC 180
DB 121 TTCTGGGGTACCGGGTGTGATTTCTGAGCCCTTCCGCAATCCCTATATTCATCAACCAAC 180
QY 181 CCCTTTGCCCGACAGCTTGGCGCTGCTGCGGCTCTGAGCTCGCCCTTCGCCCAACAGCC 240
DB 181 CCCTTTGCCCGACAGCTTGGCGCTGCTGCGGCTCTGAGCTCGCCCTTCGCCCAACAGCC 240
QY 241 CGGCCAATTGGCTCCACTTGGCGAGATCAGGCCAGCGCCCTTCGCGTCTCCGCTCGC 300
DB 241 CGGCCAATTGGCTCCACTTGGCGAGATCAGGCCAGCGCCCTTCGCGTCTCCGCTCGC 300
QY 301 CGACCTGCCACAGCGGGGTGCGCGCTGTCGCGCTGTCGCGCTGTCGCGCTGTCGCGCTCA 360
DB 301 CGACCTGCCACAGCGGGGTGCGCGCTGTCGCGCTGTCGCGCTGTCGCGCTGTCGCGCTCA 360
QY 361 CCGCTCCCGGACGTTGATTTCTCGCGGTGCAATTTCTACGCCCGCAGTATAATTTGCTACT 420
DB 361 CCGCTCCCGGACGTTGATTTCTCGCGGTGCAATTTCTACGCCCGCAGTATAATTTGCTACT 420

Db 1741 TCAACTATACACACAGGCTTTGGGCGCGTCCGGTCGCCAATTTCTGCGCGCGCGGTTTGTG 1800
QY
1801 GCTCCAGCTCGCCCTGGCTCTGCTGGAGGATACCTTTTGATATATCCGGGGGGGGCGCAC 1860
Db
1801 GCTCCAGCTCGCCCTGGCTCTGCTGGAGGATACCTTTTGATATATCCGGGGGGGGCGCAC 1860
QY
1861 ACAATTTGATGACTTCTGCGCCCTGAATCGCGCGCTTTAGGCTCCAGGGTTGTGCTTCCAG 1920
Db
1861 ACAATTTGATGACTTCTGCGCCCTGAATCGCGCGCTTTAGGCTCCAGGGTTGTGCTTCCAG 1920
QY
1921 TCACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGGTGGGTAAACCTCGGAGTTGTAG 1980
Db
1921 TCACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGGTGGGTAAACCTCGGAGTTGTAG 1980

RESULT 5

US-08-477-292-2
; Sequence 2, Application US/08477292
; Patent No. 6291641
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas
; APPLICANT: McAtee, Patrick
; APPLICANT: Yarbough, Patrice
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cathleen M. Desjardins, M.D.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: US
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,292
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/327,952
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, Allan A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: G32P5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)369-9500
; TELEFAX: (415)368-0709
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2055 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)
; INDIVIDUAL ISOLATE: ORF-2 region
US-08-477-292-2

Query Match 100.0%; Score 1980; DB 3; Length 2055;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCCCTAGGCTCTTTTGTGTTCTCTTGTTCCTGCTTCTGCTGAGTTCGCGCGCA 60
Db 1 ATGGGCCCTAGGCTCTTTTGTGTTCTCTTGTTCCTGCTTCTGCTGAGTTCGCGCGCA 60

QY 61 CCGACCGGTGAGCGGTCTGCGCGCGTCTGCGCGCGCGAGCGCGGTACCGCGGTGTGT 120
Db
61 CCGACCGGTGAGCGGTCTGCGCGCGTCTGCGCGCGCGAGCGCGGTACCGCGGTGTGT 120
QY 121 TTCTGGGGTGAGCGGGTTGATTCTCAGCCCTTTCGCAATCCCTCTATATTCATCAACCAAC 180
Db
121 TTCTGGGGTGAGCGGGTTGATTCTCAGCCCTTTCGCAATCCCTCTATATTCATCAACCAAC 180
QY 181 CCTTTGCCCCCAGAGGTTGCGCGCTGCGGTCTGACCTCGCTTCGCGCTTCGCGCAACAGCC 240
Db
181 CCTTTGCCCCCAGAGGTTGCGCGCTGCGGTCTGACCTCGCTTCGCGCTTCGCGCAACAGCC 240
QY 241 CGGCCACTTGGCTCCACCTTGGCGAGATCAGCGCCAGCGCCCTTCGCGCTTCGCGCTTCGCG 300
Db
241 CGGCCACTTGGCTCCACCTTGGCGAGATCAGCGCCAGCGCCCTTCGCGCTTCGCGCTTCGCG 300
QY 301 CGACTGCGCACAGCGGGGTGCGCGCTGAGCGGTGCGCGCTTCGCGCTTCGCGCTTCGCGCTTC 360
Db
301 CGACTGCGCACAGCGGGGTGCGCGCTGAGCGGTGCGCGCTTCGCGCTTCGCGCTTCGCGCTTC 360
QY 361 CCGTTCGGGAGCGGTGATTCTGCGGGTCAATTCAGCGCGCGAGTAAATTTGTCTACT 420
Db
361 CCGTTCGGGAGCGGTGATTCTGCGGGTCAATTCAGCGCGCGAGTAAATTTGTCTACT 420
QY 421 TCACCCCTGACATCTCTGTCGCGCTCTGCGCACTAAATTTAGTCTGATGAGCGCCCTT 480
Db
421 TCACCCCTGACATCTCTGTCGCGCTCTGCGCACTAAATTTAGTCTGATGAGCGCCCTT 480
QY 481 AATCGGCTCTGCGCTGCGCGGTACTAATCTACATTTATGCGCACAGAGSCCTCC 540
Db
481 AATCGGCTCTGCGCTGCGCGGTACTAATCTACATTTATGCGCACAGAGSCCTCC 540
QY 541 AATTATGACAGTACCGGGTTCGCGCGCTACTATCGTTACCGCGCCCTAGTGCCTAAT 600
Db
541 AATTATGACAGTACCGGGTTCGCGCGCTACTATCGTTACCGCGCCCTAGTGCCTAAT 600
QY 601 GCAGTTGGAGGTATGCTATATCAATTTCTTCTGGCTCAAAACACCAACCCCTACA 660
Db
601 GCAGTTGGAGGTATGCTATATCAATTTCTTCTGGCTCAAAACACCAACCCCTACA 660
QY 661 TCTGTGACATGAATTCATTTTCCACTGATGTCCAGATTCTTGTTCACCTGGCATA 720
Db
661 TCTGTGACATGAATTCATTTTCCACTGATGTCCAGATTCTTGTTCACCTGGCATA 720
QY 721 GCATCTGAATTTGCTATCCCAAGAGGCGCTTCACTACCGCAATCAAGTTGGGCTCG 780
Db
721 GCATCTGAATTTGCTATCCCAAGAGGCGCTTCACTACCGCAATCAAGTTGGGCTCG 780
QY 781 GTTGAGACATCTGTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Db
781 GTTGAGACATCTGTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
QY 841 CATGGCTCTCCAGTTAACTCCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 900
Db
841 CATGGCTCTCCAGTTAACTCCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 900
QY 901 GACTTTGGCTTAGAGCTTGAATTCGCAATCTCACCACCTGTAAACCAATACAGTGTG 960
Db
901 GACTTTGGCTTAGAGCTTGAATTCGCAATCTCACCACCTGTAAACCAATACAGTGTG 960
QY 961 TCCCGTTTACTCCAGACTGCTGCTCACTCCGCGCGAGGGGCGGAGCTCGGAGCTG 1020
Db
961 TCCCGTTTACTCCAGACTGCTGCTCACTCCGCGCGAGGGGCGGAGCTCGGAGCTG 1020
QY 1021 ACCAACAATCGAGCCACAGGTTTCATGAAAGATCTCCACTTTACCGGCTTTAATGGGTA 1080
Db
1021 ACCAACAATCGAGCCACAGGTTTCATGAAAGATCTCCACTTTACCGGCTTTAATGGGTA 1080
QY 1081 GGTGAAGTGGCGCGGAGTGTCTTAACATTAACCTTGTGCTGACAGCTCTCGGC 1140
Db
1081 GGTGAAGTGGCGCGGAGTGTCTTAACATTAACCTTGTGCTGACAGCTCTCGGC 1140
QY 1141 GGGCTCCGACAGAAATTAATTTGCTCGCTGGGCGCACTGTTTTATTCCCGCCGGTT 1200

Db	1141	GGGCTCCGACAGAAATTAATTTCTGTCGGCTGGCGGGCAACTGTTTTATTCCCGCCGGTT	1200
Qy	1201	GTCTCAGCCAAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAGCAG	1260
Db	1201	GTCTCAGCCAAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAGCAG	1260
Qy	1261	GATAAGGGTGTGTCTATCCCCACGATATCGATCTTTGGTGATTCGGGTGTGGTCAATTCAG	1320
Db	1261	GATAAGGGTGTGTCTATCCCCACGATATCGATCTTTGGTGATTCGGGTGTGGTCAATTCAG	1320
Qy	1321	GATTATGACAACAGCATGAGCAGGATCGGCCACCCCGTCGCTCGCCTGGCCCATCTCGCCCT	1380
Db	1321	GATTATGACAACAGCATGAGCAGGATCGGCCACCCCGTCGCTCGCCTGGCCCATCTCGCCCT	1380
Qy	1381	TTTTCCTGTTCTCCGAGCAATGATGTACTTTTGGCTGTCCTCACTCGAGCCGAGTATGAC	1440
Db	1381	TTTTCCTGTTCTCCGAGCAATGATGTACTTTGCTGTCCCTCACTCGAGCCGAGTATGAC	1440
Qy	1441	CAGTCCACTTACGGGTGCTCAACTGGCGCCCGTTTATATCTCGACACGCTGACTTTGGTG	1500
Db	1441	CAGTCCACTTACGGGTGCTCAACTGGCGCCCGTTTATATCTCGACACGCTGACTTTGGTG	1500
Qy	1501	AATGTTGCGACTGGCGCGCAGCGGTAGCCCGATCGCTTGACTGGTCCAAAGTCAACCTC	1560
Db	1501	AATGTTGCGACTGGCGCGCAGCGGTAGCCCGATCGCTTGACTGGTCCAAAGTCAACCTC	1560
Qy	1561	GACGGGGCGCCCTCCGACTGTGTGACGATATTTCCAAGACATTTCTTTGCTGCCCTT	1620
Db	1561	GACGGGGCGCCCTCCGACTGTGTGACGATATTTCCAAGACATTTCTTTGCTGCCCTT	1620
Qy	1621	CGTGGCAAGCTCTCCTTTTGGAGGGCGGCACAAACAAAGCAGGTTATCCTTATAATTAT	1680
Db	1621	CGTGGCAAGCTCTCCTTTTGGAGGGCGGCACAAACAAAGCAGGTTATCCTTATAATTAT	1680
Qy	1681	AATACTACTGCTAGTGACCAGATTCTGATTTGAAATGCTCCGCGCCATCGGGTCGCCATT	1740
Db	1681	AATACTACTGCTAGTGACCAGATTCTGATTTGAAATGCTCCGCGCCATCGGGTCGCCATT	1740
Qy	1741	TCAACTATTACACAGGCTTGGGGCCGGTCCGGTCCGCATTTCTGCGGCGCGGGTTTG	1800
Db	1741	TCAACTATTACACAGGCTTGGGGCCGGTCCGGTCCGCATTTCTGCGGCGCGGGTTTG	1800
Qy	1801	GCCTCACGCTCCGCCCTGCTGCTCGAGGATATTTTTCATTTATCCGGGGCGGGCGCAC	1860
Db	1801	GCCTCACGCTCCGCCCTGCTGCTCGAGGATATTTTTCATTTATCCGGGGCGGGCGCAC	1860
Qy	1861	ACATTTTGATGACTTCTGCCCTTGATTCGGCGGCTTTTAGCCCTCCAGGGTTGTGCTTCCAG	1920
Db	1861	ACATTTTGATGACTTCTGCCCTTGATTCGGCGGCTTTAGCCCTCCAGGGTTGTGCTTCCAG	1920
Qy	1921	TCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGGAGTTGTGTAG	1980
Db	1921	TCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGGAGTTGTGTAG	1980

RESULT 6

```

US-07-870-985A-4
; Sequence 4, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tzu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof
; TITLE OF INVENTION: Hepatitis E
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Delinger & Associates
; STREET: 350 Cambridge Avenue,
; CITY: Palo Alto

```

QY 301 CGACCTGCCACAGCGGGGCTCGCGGCTGAGCGGCTGTGGGGCTCGCCCATGACACCTCA 360
Db 301 CGACCTGCCACAGCGGGGCTCGCGGCTGAGCGGCTGTGGGGCTCGCCCATGACACCTCA 360
QY 361 CCCGTCGCCGAGGTGTGATTCTCGCGGTGCAATTCTGACGCGCGCAGTATAAATTGTCCTACT 420
Db 361 CCCGTCGCCGAGGTGTGATTCTCGCGGTGCAATTCTGACGCGCGCAGTATAAATTGTCCTACT 420
QY 421 TCACCCCTGACATCCTCTGTGGCTCTGCGCTAATTTAGTTCCTGTATGCGACCCCTT 480
Db 421 TCACCCCTGACATCCTCTGTGGCTCTGCGCTAATTTAGTTCCTGTATGCGACCCCTT 480
QY 481 AATCGGCTCTGCGGCTCGACGAGCGGTACTAATACTCACATTAATGGCCACAGAGCCCTCC 540
Db 481 AATCGGCTCTGCGGCTCGACGAGCGGTACTAATACTCACATTAATGGCCACAGAGCCCTCC 540
QY 541 AATTATGACAGTACCGGGTTGCCGCGCTACTATCCGTTACCGGCGCCCTAGTGCCTAAT 600
Db 541 AATTATGACAGTACCGGGTTGCCGCGCTACTATCCGTTACCGGCGCCCTAGTGCCTAAT 600
QY 601 GCAGTTGAGGCTATGCTATATCAATTTCTTCTGGGCTCAAAACAACCAACCCCTACA 660
Db 601 GCAGTTGAGGCTATGCTATATCAATTTCTTCTGGGCTCAAAACAACCAACCCCTACA 660
QY 661 TCTGTTGACATGAATCCCAATTACTTCCACTGATGTGAGGATTTCTGTTCAACCTGGCATA 720
Db 661 TCTGTTGACATGAATCCCAATTACTTCCACTGATGTGAGGATTTCTGTTCAACCTGGCATA 720
QY 721 GCATCTGHAATTGCTCATCCCAAGGAGCGCTTCACTACCGGAATCAAGGTTGGCGCTCG 780
Db 721 GCATCTGHAATTGCTCATCCCAAGGAGCGCTTCACTACCGGAATCAAGGTTGGCGCTCG 780
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGGAGCGACCTCCGGTCTGTGATGTATGCATA 840
Db 781 GTTGAGACATCTGGTGTGCTGAGGAGGAGCGACCTCCGGTCTGTGATGTATGCATA 840
QY 841 CATGGCTCCAGTTAACTCCTTACCAATACCCCTTATACGGGTGCCCTTGGCTTACTG 900
Db 841 CATGGCTCCAGTTAACTCCTTACCAATACCCCTTATACGGGTGCCCTTGGCTTACTG 900
QY 901 GACTTTGCTTAGAGTTGAGTTGGCAATCTCACCTGTATAACCAATACAGTGTG 960
Db 901 GACTTTGCTTAGAGTTGAGTTGGCAATCTCACCTGTATAACCAATACAGTGTG 960
QY 961 TCCGTTTACTCCAGCACTGCTGCTCACTCCGCGGAGGGCGGAGCTCGGAGCTG 1020
Db 961 TCCGTTTACTCCAGCACTGCTGCTCACTCCGCGGAGGGCGGAGCTCGGAGCTG 1020
QY 1021 ACCAACACTGCAGCCACAGGTTTCATGAAAGATCTCCACTTTACCGGCTTTAATGGGTA 1080
Db 1021 ACCAACACTGCAGCCACAGGTTTCATGAAAGATCTCCACTTTACCGGCTTTAATGGGTA 1080
QY 1081 GGTGAAGTGGCGGGGATAGCTCTAAATTAACCTTGTGATGACAGCTCTCGGC 1140
Db 1081 GGTGAAGTGGCGGGGATAGCTCTAAATTAACCTTGTGATGACAGCTCTCGGC 1140
QY 1141 GGGCTCCGACAGAAATAATTTCGTCGCTGCGGCAACTGTTTATTCGCGCCGCT 1200
Db 1141 GGGCTCCGACAGAAATAATTTCGTCGCTGCGGCAACTGTTTATTCGCGCCGCT 1200
QY 1201 GTCTCAGCAATGGCGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAGAG 1260
Db 1201 GTCTCAGCAATGGCGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAGAG 1260
QY 1261 GATAAGGTTGTGCTATCCCGCAGATATCGATTTGTTGATTCGCGTGTGATTCAG 1320
Db 1261 GATAAGGTTGTGCTATCCCGCAGATATCGATTTGTTGATTCGCGTGTGATTCAG 1320
QY 1321 GATTATGACCAACAGCATGAGCAGATCGGCCACCCCGTCCGCTCGCCATCTCGGCT 1380
Db 1321 GATTATGACCAACAGCATGAGCAGATCGGCCACCCCGTCCGCTCGCCATCTCGGCT 1380
QY 1381 TTTTCTGTCTCCGAGCAATGATGTACTTTTGGCTGTCCCTCACTGCGAGCGGATGAC 1440

Db 1381 TTTTCTGTCTCCGAGCAATGATGTACTTTTGGCTGTCCCTCACTGCGAGCCGATGAC 1440
QY 1441 CAGTCCACTTACGGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCGTGACTTTGGTG 1500
Db 1441 CAGTCCACTTACGGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCGTGACTTTGGTG 1500
QY 1501 AATGTTGCGACTGGCGGCGAGCGGCTAGCCGATCGCTTGAATTCGCAAGTCAACCTC 1560
Db 1501 AATGTTGCGACTGGCGGCGAGCGGCTAGCCGATCGCTTGAATTCGCAAGTCAACCTC 1560
QY 1561 GACGGGCGGCGCTCCCGACTGTTGAGCAATATTCGAAGACATTTCTTGTGCTCCCGCTT 1620
Db 1561 GACGGGCGGCGCTCCCGACTGTTGAGCAATATTCGAAGACATTTCTTGTGCTCCCGCTT 1620
QY 1621 CGTGCAAGCTCTCCTTTTGGGAGCGCGGCAACAAGAGGTTATCTTATAAATTAAT 1680
Db 1621 CGTGCAAGCTCTCCTTTTGGGAGCGCGGCAACAAGAGGTTATCTTATAAATTAAT 1680
QY 1681 AATACTACTGCTAGTGAACAGATTTCTGATTGAAATGCTCGGCGCATCGGGTGGCCATT 1740
Db 1681 AATACTACTGCTAGTGAACAGATTTCTGATTGAAATGCTCGGCGCATCGGGTGGCCATT 1740
QY 1741 TCAACCTATACACAGGCTTTGGGCGCGGTCGGTTCGCCATTTCTGCGGCGCGGTTTG 1800
Db 1741 TCAACCTATACACAGGCTTTGGGCGCGGTCGGTTCGCCATTTCTGCGGCGCGGTTTG 1800
QY 1801 GCTCAGCTCCGCGCTGCTGCTGAGAGATATCTTGTGATTAATTCGCGGCGGCGCAC 1860
Db 1801 GCTCAGCTCCGCGCTGCTGCTGAGAGATATCTTGTGATTAATTCGCGGCGGCGCAC 1860
QY 1861 ACATTTGATGACTTCTGCCCTGAATGCGCGCTTAGCCCTCCAGGTTGTCTTCCAG 1920
Db 1861 ACATTTGATGACTTCTGCCCTGAATGCGCGCTTAGCCCTCCAGGTTGTCTTCCAG 1920
QY 1921 TCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAAGTGGGTAAAGTGGGTGG 1980
Db 1921 TCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAAGTGGGTGGGTGG 1980

RESULT 7

US-08-259-148A-2
; Sequence 2, Application US/08259148A
; Patent No. 5741490
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; APPLICANT: Yarbough, Patrice D.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,148A
; FILING DATE: 13-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: MEXICO, FIGURE 7
; US-08-259-148A-2

Query Match 100.0%; Score 1980; DB 1; Length 2100;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCCCTAGCGCTCTTTGCTGTTGTTCTCTGTTCTGCTATGTTGCCGGCCA 60
DB |||||||
QY 46 ATGCGCCCTAGCGCTCTTTGCTGTTGTTCTCTGTTCTGCTATGTTGCCGGCCA 105
DB |||||||
QY 61 CCGACCGGTACCGCTCTGCGCGCCGCTGCGCGCGCGCGCGGTACCGCGGTGT 120
DB |||||||
QY 106 CCGACCGGTACCGCTCTGCGCGCCGCTGCGCGCGCGCGCGGTACCGCGGTGT 165
DB |||||||
QY 121 TTCTGGGTGACCGGTGATTTCTAGCCCTTTCGCAATCCCTATATTCATCAACCAAC 180
DB |||||||
QY 166 TTCTGGGTGACCGGTGATTTCTAGCCCTTTCGCAATCCCTATATTCATCAACCAAC 225
DB |||||||
QY 181 CCCTTTGCCCGACAGCTTGCCTGCGTCTGCGGTCTGGACCTTGCCTTCGCCAACCGAC 240
DB |||||||
QY 226 CCCTTTGCCCGACAGCTTGCCTGCGTCTGCGGTCTGGACCTTGCCTTCGCCAACCGAC 285
DB |||||||
QY 241 CGGCCACTTGGTCCACTTTGGCGAGATCAGGCCCGCGCCCTCGCTCGCTCCCGTCGC 300
DB |||||||
QY 286 CGGCCACTTGGTCCACTTTGGCGAGATCAGGCCCGCGCCCTCGCTCGCTCCCGTCGC 345
DB |||||||
QY 301 CGACCTGCCACAGCGGGGCTGGGCGCTGAGCGGTGTTGGCGCTGCCCATGACACTCA 360
DB |||||||
QY 346 CGACCTGCCACAGCGGGGCTGGGCGCTGAGCGGTGTTGGCGCTGCCCATGACACTCA 405
DB |||||||
QY 361 CCGCTCCCGGAGCTGATTTCTCGCGGTGCAATTTCTACGCGCGCGAGTAAATTTGCTACT 420
DB |||||||
QY 406 CCGCTCCCGGAGCTGATTTCTCGCGGTGCAATTTCTACGCGCGCGAGTAAATTTGCTACT 465
DB |||||||
QY 421 TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTTAGTCTGTATGAGCCGCCCTT 480
DB |||||||
QY 466 TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTTAGTCTGTATGAGCCGCCCTT 525
DB |||||||
QY 481 AATCCGCTCTCGCGCTGAGGACGGTACTAATCTCATATTATGSCACAGAGGCTCC 540
DB |||||||
QY 526 AATCCGCTCTCGCGCTGAGGACGGTACTAATCTCATATTATGSCACAGAGGCTCC 585
DB |||||||

; 541 AATTATGCACGTACCGGTTGCCCGCGCTACTATCCGTTACCGGCCCTAGTGCCCTAAT 600
; |||||||
; 586 AATTATGCACGTACCGGTTGCCCGCGCTACTATCCGTTACCGGCCCTAGTGCCCTAAT 645
; |||||||
; 601 GCAGTTGAGGCTATGCTATATCCATTTCTTTCTGGCTCAAAACCAACCAACCCCTACA 660
; |||||||
; 646 GCAGTTGAGGCTATGCTATATCCATTTCTTTCTGGCTCAAAACCAACCAACCCCTACA 705
; |||||||
; 661 TCTGTTGACATGAATTCATTTCCACTGATGTCAGGATTTCTGTTCAACTGGCATA 720
; |||||||
; 706 TCTGTTGACATGAATTTCCATTTCCACTGATGTCAGGATTTCTGTTCAACTGGCATA 765
; |||||||
; 721 GCATCTGAATTTGCTCATCCCAAGCAGCGCTTCACTACCGCAATCAAGTTGGCGCTCG 780
; |||||||
; 766 GCATCTGAATTTGCTCATCCCAAGCAGCGCTTCACTACCGCAATCAAGTTGGCGCTCG 825
; |||||||
; 781 GTTGAGACATCTGTTGTTGCTGAGGAGGAGCCACTCTCGGCTTGTGTCATGTTATGCATA 840
; |||||||
; 826 GTTGAGACATCTGTTGTTGCTGAGGAGGAGCCACTCTCGGCTTGTGTCATGTTATGCATA 885
; |||||||
; 841 CATGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 900
; |||||||
; 886 CATGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 945
; |||||||
; 901 GACTTTGCTTAGAGCTTAGGTTTCGCAATCTCACCACTGTACACCAATACAGTGTG 960
; |||||||
; 946 GACTTTGCTTAGAGCTTAGGTTTCGCAATCTCACCACTGTACACCAATACAGTGTG 1005
; |||||||
; 961 TCCGTTACTCCAGCACTGCTGTCATCTCGCGCGGAGGCGCGAGCTGCGGAGCTG 1020
; |||||||
; 1005 TCCGTTACTCCAGCACTGCTGTCATCTCGCGCGGAGGCGCGAGCTGCGGAGCTG 1065
; |||||||
; 1021 ACCCAACTGCAGCCACCAAGTTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080
; |||||||
; 1066 ACCCAACTGCAGCCACCAAGTTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA 1125
; |||||||
; 1081 GGTGAAGTCCGCGCGGATAGCTCTAACTTAATTAATTAATTAATTAATTAATTAATTAAT 1140
; |||||||
; 1126 GGTGAAGTCCGCGCGGATAGCTCTAACTTAATTAATTAATTAATTAATTAATTAATTAAT 1185
; |||||||
; 1141 GGGTCCCGACAGAAATTAATTTGCTGCGTGGGCGCAACTGTTTATTTATTTATTTATTTATTT 1200
; |||||||
; 1186 GGGCTCCCGACAGAAATTAATTTGCTGCGTGGGCGCAACTGTTTATTTATTTATTTATTTATTT 1245
; |||||||
; 1201 GTCTCAGCAATGGCGAGCCAACTGTAAGCTCTATACATCAGTGGAGAAATGCTCAGCAG 1260
; |||||||
; 1246 GTCTCAGCAATGGCGAGCCAACTGTAAGCTCTATACATCAGTGGAGAAATGCTCAGCAG 1305
; |||||||
; 1261 GATAAGGGTGTGCTATCCCCCAGCATATCGATCTTTGGTGATTCGCGTGTGGTCAATTCAG 1320
; |||||||
; 1306 GATAAGGGTGTGCTATCCCCCAGCATATCGATCTTTGGTGATTCGCGTGTGGTCAATTCAG 1365
; |||||||
; 1321 GATTATGACCAACAGCATGAGCAGGATCGGCCCAACCGCTGCGCTGCGGCATCTCGGCT 1380
; |||||||
; 1366 GATTATGACCAACAGCATGAGCAGGATCGGCCCAACCGCTGCGCTGCGGCATCTCGGCT 1425
; |||||||
; 1381 TTTTCTGTTCTCCGAGCAATGATGATCTTTGGCTGTCCCTCACTGACGCGCAGATGAC 1440
; |||||||
; 1426 TTTTCTGTTCTCCGAGCAATGATGATCTTTGGCTGTCCCTCACTGACGCGCAGATGAC 1485
; |||||||
; 1441 CAGTCCACTTACGGGTTCGTCAACTGGCCCGGTTTATATCTCGGACAGCGTGTGTTGGTG 1500
; |||||||
; 1486 CAGTCCACTTACGGGTTCGTCAACTGGCCCGGTTTATATCTCGGACAGCGTGTGTTGGTG 1545
; |||||||
; 1501 AATGTTGGACTGGCGCGCAGCGCTAGCCCATCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
; |||||||
; 1546 AATGTTGGACTGGCGCGCAGCGCTAGCCCATCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1605
; |||||||
; 1561 GACGGCGCGCCCTCCCGACTGTTGAGCAATATTCAGACATATTTCTTTGTGCTCCCGCTT 1620
; |||||||
; 1606 GACGGCGCGCCCTCCCGACTGTTGAGCAATATTCAGACATATTTCTTTGTGCTCCCGCTT 1665
; |||||||
```


Qy	721	GCATCTGAATTGGTCA	TCCCAAGCAGCGCTTC	ACTACCGCAATCAAGT	TGGCGCTCG	780
Db	766	GCATCTGAATTGGTCA	TCCCAAGCAGCGCTTC	ACTACCGCAATCAAGT	TGGCGCTCG	825
Qy	781	GTTTGAGACATCTGGT	TTGCTGAGGAGGAGC	CACTCCGGTCTCTG	TCATGTTATTCGATA	840
Db	826	GTTTGAGACATCTGGT	TTGCTGAGGAGGAGC	CACTCCGGTCTCTG	TCATGTTATTCGATA	885
Qy	841	CATGGCTCTCCAGTTAA	CTCCTATACCAATAC	CCCCCTTATACCGGT	GCCTTGCTTACTCTG	900
Db	886	CATGGCTCTCCAGTTAA	CTCCTATACCAATAC	CCCCCTTATACCGGT	GCCTTGCTTACTCTG	945
Qy	901	GACTTTGCGCTTAGAG	CTTGAGTTTCGCAATCT	CACACCTGTAAACCA	CAATACACCTGTG	960
Db	946	GACTTTGCGCTTAGAG	CTTGAGTTTCGCAATCT	CACACCTGTAAACCA	CAATACACCTGTG	1005
Qy	961	TCCGGTTACTCCAGCA	CTGCTCGTCACTCCG	CCGAGGGCCGACGGGA	CTCGCGAGCTG	1020
Db	1006	TCCGGTTACTCCAGCA	CTGCTCGTCACTCCG	CCGAGGGCCGACGGGA	CTCGCGAGCTG	1065
Qy	1021	ACCAAACTGAGCGCA	CAGGTTTCATGAAAGAT	CTTCACTTTACCGGC	CTTAAATGGGTTA	1080
Db	1066	ACCAAACTGAGCGCA	CAGGTTTCATGAAAGAT	CTTCACTTTACCGGC	CTTAAATGGGTTA	1125
Qy	1081	GGTGAAGTCGCGCGG	GATAGCTTAACATTA	CTTAACCTTGCTGAC	ACGCTCTCTCGGC	1140
Db	1126	GGTGAAGTCGCGCGG	GATAGCTTAACATTA	CTTAACCTTGCTGAC	ACGCTCTCTCGGC	1185
Qy	1141	GGGCTCCGACAGAA	TAAATTCGTGCGCT	GCGGGCAACTGTTT	TATATCCCGCCGGTT	1200
Db	1186	GGGCTCCGACAGAA	TAAATTCGTGCGCT	GCGGGCAACTGTTT	TATATCCCGCCGGTT	1245
Qy	1201	GTCTCAGCAATGCGAG	CCAAACGCTGAAGCT	CTATACATCAGTGG	AGAAATGCTCAGCAG	1260
Db	1246	GTCTCAGCAATGCGAG	CCAAACGCTGAAGCT	CTATACATCAGTGG	AGAAATGCTCAGCAG	1305
Qy	1261	GATAAGGGTGTGTCT	ATCCCCCAGATATCG	ATCTTGGTGATTCCG	TGTGGTCAATTCAG	1320
Db	1306	GATAAGGGTGTGTCT	ATCCCCCAGATATCG	ATCTTGGTGATTCCG	TGTGGTCAATTCAG	1365
Qy	1321	GATTATGACAACAGAT	GACAGATGCGCCAC	CGCGCTCGCTCGCC	ATCTCGGCGCT	1380
Db	1366	GATTATGACAACAGAT	GACAGATGCGCCAC	CGCGCTCGCTCGCC	ATCTCGGCGCT	1425
Qy	1381	TTTTCTGTCTCCGAG	CAAAATGATGTACT	TTTGGTGTCCCTCA	CTGCGAGCGGAGTATGAC	1440
Db	1426	TTTTCTGTCTCCGAG	CAAAATGATGTACT	TTTGGTGTCCCTCA	CTGCGAGCGGAGTATGAC	1485
Qy	1441	CAGTCCACTTACGGGT	CGTCAACTGGCCCGG	TTTATATCTCGGA	CACGGTGACTTTGGTG	1500
Db	1486	CAGTCCACTTACGGGT	CGTCAACTGGCCCGG	TTTATATCTCGGA	CACGGTGACTTTGGTG	1545
Qy	1501	AATGTTGCGACTGGG	CGCAGCGGTAGCCG	ATGCTTGACTCGGT	CCAAAGTCACCCCTC	1560
Db	1546	AATGTTGCGACTGGG	CGCAGCGGTAGCCG	ATGCTTGACTCGGT	CCAAAGTCACCCCTC	1605
Qy	1561	GACGGCGGCCCTCTC	CCCGACTGTTGAGCA	ATATTTCAAGACAT	TCTTTGTGCTCCCCCTT	1620
Db	1606	GACGGCGGCCCTCTC	CCCGACTGTTGAGCA	ATATTTCAAGACAT	TCTTTGTGCTCCCCCTT	1665
Qy	1621	CGTGGCAAGCTCTCC	TTTTTGGAGGCGGCA	CAAAAGCAGGTAT	CTTATTAATAT	1680
Db	1666	CGTGGCAAGCTCTCC	TTTTTGGAGGCGGCA	CAAAAGCAGGTAT	CTTATTAATAT	1725
Qy	1681	AATACTACTGTAGTGA	CCAGATTCGTGAAAT	TGCTCGGCCCAT	TCGGGTCCGCATT	1740
Db	1726	AATACTACTGTAGTGA	CCAGATTCGTGAAAT	TGCTCGGCCCAT	TCGGGTCCGCATT	1785
Qy	1741	TCAACTATATACACAG	GGTTGGGCGCGGT	TCGGTCCGCAAT	TCTGCGCGCGGTTTGTG	1800
Db	1786	TCAACTATATACACAG	GGTTGGGCGCGGT	TCGGTCCGCAAT	TCTGCGCGCGGTTTGTG	1845
Qy	1801	GCTTCACGCTCCGCC	CTGGCTCTGTGGAG	GATATTTTGTATAT	CTTCGGGCGGCGGCAC	1860

```

Db      1846  GCTCCACGCTCGGCCCTGGCTCTGCTGGAGGATACTTTTGATTATCCGGGGCGGGCGCAC 1905
Qy      1861  ACATTTTGATGACTTCTGCCCCCTGAATCGCCGCTTTTAGCGCTCCACAGGTTTGCTGTTTCCAG 1920
Db      1906  ACAITTTGATGACTTCTGCCCCCTGAATCGCCGCTTTTAGCGCTCCACAGGTTTGCTGTTTCCAG 1965
Qy      1921  TCAACTGTCGCTGAGCTCCAGCGCCCTTAAAGTTTAAAGTGGGTGGAATAAACTCCGGGAGCTTCTGAG 1980
Db      1966  TCAACTGTCGCTGAGCTCCAGCGCCCTTAAAGTTTAAAGTGGGTGGAATAAACTCCGGGAGTTGTAG 2025

RESULT 9
US-07-876-941A-2
; Sequence 2, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876,941A
; FILING DATE: 01-MAY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.33
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO

```


APPLICANT: Reyes, Gregory R.
APPLICANT: Bradley, Daniel W.
APPLICANT: Wu, Jr-Shin
APPLICANT: Purdy, Michael A.
APPLICANT: Tam, Albert W.
APPLICANT: Krawczynski, Krzysztof Z.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/870,985A
FILING DATE: 20-APRIL-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0093.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2100 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: MEXICO, FIGURE 7
US-07-870-985A-2

Query Match 100.0%; Score 1980; DB 4; Length 2100;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCCCTAGGCGCTCTTTTGTGTTGTTCTTCTGCTATGTTGCGCTATGTTGCGCGGCCA 60
DB |||||||
QY 61 CCGACCGGTCAGCGCTCGTCG 120
DB |||||||
QY 106 CCGACCGGTCAGCGCTCGTCG 165
DB |||||||

121 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
DB |||||||
166 TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 225
DB |||||||
181 CCCTTTGCCCGGAGAGCGTTGCGCGCTGCGCGGTCTGACCTTCGCGCTTCGCCAACCGGC 240
DB |||||||
226 CCCTTTGCCCGGAGAGCGTTGCGCGCTGCGCGGTCTGACCTTCGCGCTTCGCCAACCGGC 285
DB |||||||
241 CGGCGCACTTGGCTCCACTTTGGCGAGATCAGGCGCCGAGCGCCCTCGCTCCCTCCCGTCG 300
DB |||||||
286 CGGCGCACTTGGCTCCACTTTGGCGAGATCAGGCGCCGAGCGCCCTCGCTCCCTCCCGTCG 345
DB |||||||
301 CGACCTGCGCACAGCGCGGCTGCGCGGCTGACGGCTGTGGCGCTGCGCCATGACACTCA 360
DB |||||||
346 CGACCTGCGCACAGCGCGGCTGCGCGGCTGACGGCTGTGGCGCTGCGCCATGACACTCA 405
DB |||||||
361 CCGGTCGCGGAGCGTTGATTCTCGCGGTGCAATTTCTACGCGCGCAGTATATTTGTTCTACT 420
DB |||||||
406 CCGGTCGCGGAGCGTTGATTCTCGCGGTGCAATTTCTACGCGCGCAGTATATTTGTTCTACT 465
DB |||||||
421 TCACCCCTGACATCTCTCTGTGGCTCTGGCACTAAATTTAGTCTCTGTATGACGCCCCCTT 480
DB |||||||
466 TCACCCCTGACATCTCTCTGTGGCTCTGGCACTAAATTTAGTCTCTGTATGACGCCCCCTT 525
DB |||||||
481 AATCCGCTCTCGCGCTGCGAGCGGTACTAATATCAATATGAGCCACAGAGGCTCC 540
DB |||||||
526 AATCCGCTCTCGCGCTGCGAGCGGTACTAATATCAATATGAGCCACAGAGGCTCC 585
DB |||||||
541 AATTATGACAGTACCGGGTTGCGCGGCTACTATCCGTTACCGGCCCTAGTGCTTAAT 600
DB |||||||
586 AATTATGACAGTACCGGGTTGCGCGGCTACTATCCGTTACCGGCCCTAGTGCTTAAT 645
DB |||||||
601 GCAGTTGGAGGCTATGCTATATCCATTTCTTGGCTTCAAAACCAACCAACCCCTACA 660
DB |||||||
646 GCAGTTGGAGGCTATGCTATATCCATTTCTTGGCTTCAAAACCAACCAACCCCTACA 705
DB |||||||
661 TCTGTTGACATGAATTCATTTCCACTGATGTCCAGGATTTCTTGTTCAACTGGCATA 720
DB |||||||
706 TCTGTTGACATGAATTCATTTCCACTGATGTCCAGGATTTCTTGTTCAACTGGCATA 765
DB |||||||
721 GCATCTGAATTTGGTCAATCCAGCGAGCGCTTCACTACGCAATCAAGGTTGGCGCTCG 780
DB |||||||
766 GCATCTGAATTTGGTCAATCCAGCGAGCGCTTCACTACGCAATCAAGGTTGGCGCTCG 825
DB |||||||
781 GTTGAGACATCTGGTGTCTGCTGAGAGGAGGACACCTCCGCTCTTCTCATGTTATGCATA 840
DB |||||||
826 GTTGAGACATCTGGTGTCTGCTGAGAGGAGGACACCTCCGCTCTTCTCATGTTATGCATA 885
DB |||||||
841 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 900
DB |||||||
886 CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 945
DB |||||||
901 GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACCACTGTAAACCAATACACCTGTG 960
DB |||||||
946 GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACCACTGTAAACCAATACACCTGTG 1005
DB |||||||
961 TCCCGTTTACTCCAGCACTCTCGTCACTCCGCGCGGAGGCGCGACGCGGAGCTGCGAGCTG 1020
DB |||||||
1006 TCCCGTTTACTCCAGCACTCTCGTCACTCCGCGCGGAGGCGCGACGCGGAGCTGCGAGCTG 1065
DB |||||||
1021 ACCAACAATGCGAGCCACAGGTTTCATGAAGAATCTCCACTTTACCGGCTTAAATGGGGTA 1080
DB |||||||
1066 ACCAACAATGCGAGCCACAGGTTTCATGAAGAATCTCCACTTTACCGGCTTAAATGGGGTA 1125
DB |||||||
1081 GGTGAAGTGGCGCGGAGTAGCTTAACATTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1140
DB |||||||
1126 GGTGAAGTGGCGCGGAGTAGCTTAACATTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1185
DB |||||||
1141 GGGCTCCCCACAGAAATTAATTTGCTCGGTGGCGGCGCAACTGTTTAACTTAACTTAACTTAA 1200
DB |||||||
1186 GGGCTCCCCACAGAAATTAATTTGCTCGGTGGCGGCGCAACTGTTTAACTTAACTTAACTTAA 1245
DB |||||||
1201 GTCTCAGCCCAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAGCAG 1260
DB |||||||

Qy	301	CGACCTGCCACAGCCGGGCTGCGCGCTGACGGCTGTGGGGCTGCCCAATGACACTCA	360
Db	5417	CGACCTGCCACAGCTCGGGCTGCGCGCTGACGGCTGTGGCGCTGCCATGACACCTCA	5476
Qy	361	CCCGTCCGGACGTTGATCTCGGGGTGCAATCTACGCCCGCAGTATAATTTGTCTACT	420
Db	5477	CCCGTCCGGACGTTGATCTCGGGGTGCAATCTACGCCCGCAGTATAATTTGTCTACT	5536
Qy	421	TCACCCCTGACATCCTCTGTGGCTCTGGCACTAAATTTAGTCTCTGTATGAGCCCCCTT	480
Db	5537	TCACCCCTGACATCCTCTGTGGCTCTGGCACTAAATTTAGTCTCTGTATGAGCCCCCTT	5596
Qy	481	AATCCGCTCTGCGCTGCGAGGACGGTACTAAATCTCACTATATGGCCACAGAGGCTCC	540
Db	5597	AATCCGCTCTGCGCTGCGAGGACGGTACTAAATCTCACTATATGGCCACAGAGGCTCC	5656
Qy	541	AAATATGCACAGTACCGGTTGCCCGCTACTATCCGTTACCGGCCCTAGTGCCTAAT	600
Db	5657	AAATATGCACAGTACCGGTTGCCCGCTACTATCCGTTACCGGCCCTAGTGCCTAAT	5716
Qy	601	GCAGTGTGAGGCTATGCTATATCAATTTCTTTCTGGCTCAAAACAACCAACCCCTACA	660
Db	5717	GCAGTGTGAGGCTATGCTATATCAATTTCTTTCTGGCTCAAAACAACCAACCCCTACA	5776
Qy	661	TCTGTTGACATGAATTCATTTACTTTCACATGTCAGAGATCTTGTTCACCTGGGATA	720
Db	5777	TCTGTTGACATGAATTCATTTACTTTCACATGTCAGAGATCTTGTTCACCTGGGATA	5836
Qy	721	GCATCTGAATTTGTCATCCCAAGCGAGCGCTTCACTACCGCAATCAAGTTTGGCGCTG	780
Db	5837	GCATCTGAATTTGTCATCCCAAGCGAGCGCTTCACTACCGCAATCAAGTTTGGCGCTG	5896
Qy	781	GTTGAGACATCTGCTGTTGCTGAGGAGGAGCCACCTCCGCTCTTGTTCATGTTATGCATA	840
Db	5897	GTTGAGACATCTGCTGTTGCTGAGGAGGAGCCACCTCCGCTCTTGTTCATGTTATGCATA	5956
Qy	841	CATGGCTCTCCAGTTAACTCTTAFACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG	900
Db	5957	CATGGCTCTCCAGTTAACTCTTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG	6016
Qy	901	GACTTTGCTTTAGAGCTTGAGTTTCGCAATCTACCACTGTACACCAATACACGNGTG	960
Db	6017	GACTTTGCTTTAGAGCTTGAGTTTCGCAATCTACCACTGTACACCAATACACGNGTG	6076
Qy	961	TCCCGTTACTCCAGCACTGCTCGTCACTCCGCCCGAGGGGCCGAGCACTCGGAGCTG	1020
Db	6077	TCCCGTTACTCCAGCACTGCTCGTCACTCCGCCCGAGGGGCCGAGCACTCGGAGCTG	6136
Qy	1021	ACCACAACTGACGCCACAGGTTCAATGAAGATCTCACTTTACCGGCTTAAATGGGGTA	1080
Db	6137	ACCACAACTGACGCCACAGGTTCAATGAAGATCTCACTTTACCGGCTTAAATGGGGTA	6196
Qy	1081	GGTGAAGTCCGCGCGGATAGCTCTAACATTTAATTTTACCTTGCTGACAGCTCTCGGC	1140
Db	6197	GGTGAAGTCCGCGCGGATAGCTCTAACATTTAATTTTACCTTGCTGACAGCTCTCGGC	6256
Qy	1141	GGGCTCCCGACAGAAATTAATTTGCTGGCTCGCGGGCAACTGTTTTTATTCGCGCCGGTT	1200
Db	6257	GGGCTCCCGACAGAAATTAATTTGCTGGCTCGCGGGCAACTGTTTTTATTCGCGCCGGTT	6316
Qy	1201	GTCTCAGCAATGGCGAGCCAAACCGTGAAGCTCTATACATCAAGTGGAGAAATGCTCAGAG	1260
Db	6317	GTCTCAGCAATGGCGAGCCAAACCGTGAAGCTCTATACATCAAGTGGAGAAATGCTCAGAG	6376
Qy	1261	GATAAGGTTGCTATCCCGACAGATTCGATCTGGTGATTTCGGTTGCTCATTCAG	1320
Db	6377	GATAAGGTTGCTATCCCGACAGATTCGATCTGGTGATTTCGGTTGCTCATTCAG	6436
Qy	1321	GATTATGACAAACCAAGATGAGCAGGATCGGCCAACCCCGCTCGCTGGGCCATCTCGGCT	1380
Db	6437	GATTATGACAAACCAAGATGAGCAGGATCGGCCAACCCCGCTCGCTGGGCCATCTCGGCT	6496
Qy	1381	TTTTCTGTTCTCCGAGAAATGATGATCTTTTGGCTGTCCTTCACTGCGAGGAGTATGAC	1440

Db	6497	TTTTCTCTTCTCGAGAAATGATGTACTTTTGCTGTCCCTCACTCGACCGAGTATGAC	6556
Qy	1441	CAGTCCACTTACCGGTCGTCAACTGGCCCGGTTTATATCTCGACACGCTGACTTTGGTG	1500
Db	6557	CAGTCCACTTACCGGTCGTCAACTGGCCCGGTTTATATCTCGACACGCTGACTTTGGTG	6616
Qy	1501	AATGTTCCGACTGGCGGCGAGCGCGTAGCCCGATCGCTTGACTGGTCCAAAGTCAAGCCCTC	1560
Db	6617	AATGTTCCGACTGGCGGCGAGCGCGTAGCCCGATCGCTTGACTGGTCCAAAGTCAAGCCCTC	6676
Qy	1561	GACGGGCGGCCCTCCCGACTGTGTGAGCAATATTCAAAGACATTTCTTTGTGCTCCGCCCTT	1620
Db	6677	GACGGGCGGCCCTCCCGACTGTGTGAGCAATATTCAAAGACATTTCTTTGTGCTCCGCCCTT	6736
Qy	1621	CGTGGCAAGCTCTCTTTTGGGAGGCCGGGACAAACAAAGCAGGTTATCCTTATAATTAT	1680
Db	6737	CGTGGCAAGCTCTCTTTTGGGAGGCCGGGACAAACAAAGCAGGTTATCCTTATAATTAT	6796
Qy	1681	AATACTACTGCTAGTGACAGATTTCTGATTGAAATGCTCCCGGCCATCGGGTCGCAATT	1740
Db	6797	AATACTACTGCTAGTGACAGATTTCTGATTGAAATGCTCCCGGCCATCGGGTCGCAATT	6856
Qy	1741	TCAACCTATACCAACAGGCTTGGGGCGGTCGCGTCCGCAATTTCTGCGCGCGGTTTTG	1800
Db	6857	TCAACCTATACCAACAGGCTTGGGGCGGTCGCGTCCGCAATTTCTGCGCGCGGTTTTG	6916
Qy	1801	GCTCCAGCTCCGCCCTGCTCTGCTGGAGATACATTTTGATTTATCCGGGCGGCGCAC	1860
Db	6917	GCTCCAGCTCCGCCCTGCTCTGCTGGAGATACATTTTGATTTATCCGGGCGGCGCAC	6976
Qy	1861	ACATTTGATGACTTCTGCCCTGAATGCGCGCTTTTAGCGCTCCAGGGTTGTGCTTTCCAG	1920
Db	6977	ACATTTGATGACTTCTGCCCTGAATGCGCGCTTTTAGCGCTCCAGGGTTGTGCTTTCCAG	7036
Qy	1921	TCAACTGTCGTGAGCTCCAGCGCCTTAAAGTTAAGTGGGTAAAACCTCGGAGTTGTAG	1980
Db	7037	TCAACTGTCGTGAGCTCCAGCGCCTTAAAGTTAAGTGGGTAAAACCTCGGAGTTGTAG	7096

RESULT 12

US-09-128-275A-10

; Sequence 10, Application US/09128275A

; Patent No. 6229005

; GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory R

; APPLICANT: Yarbough, Patrice O

; APPLICANT: Bradley, Daniel W

; APPLICANT: Krawczynski, Krzysztof Z

; APPLICANT: Tam, Albert

; APPLICANT: Fry, Kirk E

; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted

; TITLE OF INVENTION: No. 6229005-A/No. 6229005-B Hepatitis Viral Agent

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Denlinger & Associates

; STREET: 350 Cambridge Avenue, Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/128,275A

; FILING DATE: 03-AUG-1998

; PRIOR APPLICATION DATA:

; PRIOR APPLICATION NUMBER: US 08/279,823

; FILING DATE: 25-JUL-1994

; PRIOR APPLICATION DATA:

QY	61	CCGACCGGCTCAGCCGCTCTGGCGCGCGTCTGGCGCGCGAGCGCGGTACCGCGCGTGGT	120	Db	1141	GGCGGGCTCCCGACAGAAATTAATTTCTGTCGGCTGGCGGCAACTGTTTTATTTCCGCGCG	1200
Db	61	CCGACCGGCTCAGCCGCTCTGGCGCGCGTCTGGCGCGCGAGCGCGGTACCGCGCGTGGT	120	QY	1198	GTTCGTCTCAGCCAAATGCGGACCAACCGTGAAGCTCTATACATCAGTGAAGAAATGCTCAG	1257
QY	121	TTCTGGGGTGACCGGGTGAATTTCTCAGCCCTTCGCAATCCCTCATATATTCATCAACCAAC	180	Db	1201	GTTCGTCTCAGCCAAATGCGGACCAACCGTGAAGCTCTATACATCAGTGAAGAAATGCTCAG	1260
Db	121	TTCTGGGGTGACCGGGTGAATTTCTCAGCCCTTCGCAATCCCTCATATATTCATCAACCAAC	180	QY	1258	CAGGATAAGGGTGTTCCTATCCCCACGATATCGATCTTTGGTGATTTCCGCTGTGGTCAAT	1317
QY	181	CCCTTTGCCCCAGAGATTGCCGTGCTGGGTCTGGAGCTCGGCTCGGCCAAACGAGCC	240	Db	1261	CAGGATAAGGGTGTTCCTATCCCCACGATATCGATCTTTGGTGATTTCCGCTGTGGTCAAT	1320
Db	181	CCCTTTGCCCCAGAGATTGCCGTGCTGGGTCTGGAGCTCGGCTCGGCCAAACGAGCC	240	QY	1318	CAGGATAATGACCAACGAGCATGAGGATCGGCCCAACCCCGTGGCTCGGCCAATCTCGG	1377
QY	241	CGGCGACTTGGCTCCACATTTGGGAGATCAGGGCCAGCGCCCTCGCTCCGCTCGCGTGC	300	Db	1321	CAGGATAATGACCAACGAGCATGAGGATCGGCCCAACCCCGTGGCTCGGCCAATCTCGG	1380
Db	241	CGGCGACTTGGCTCCACATTTGGGAGATCAGGGCCAGCGCCCTCGCTCCGCTCGCGTGC	300	QY	1378	CTTTTTCTGTGTTCTCGGACCAATGATGTACTTTGGCTCTCCTCACTCGAGCCGAGTAT	1437
QY	301	CGACCTGCCACAGCGCGGGCTCGGGCGCTGACGGCTGTGGCGCTGCCCATGACACCTCA	360	Db	1381	CTTTTTCTGTGTTCTCGGACCAATGATGTACTTTGGCTCTCCTCACTCGAGCCGAGTAT	1440
Db	301	CGACCTGCCACAGCGCGGGCTCGGGCGCTGACGGCTGTGGCGCTGCCCATGACACCTCA	360	QY	1438	GACCACTCCACATTAAGGGTTCGTCAACTGGCCCGGTTTATATCTCGGACAGCGTGAATTTG	1497
QY	361	CCGCTCCCGGAGCTTGAATTTCTCGGGTGCAATTTCTACGGCGCCAGTATATTTGTCTACT	420	Db	1441	GACCACTCCACATTAAGGGTTCGTCAACTGGCCCGGTTTATATCTCGGACAGCGTGAATTTG	1500
Db	361	CCGCTCCCGGAGCTTGAATTTCTCGGGTGCAATTTCTACGGCGCCAGTATATTTGTCTACT	420	QY	1498	GTGAATGTTGCGACTGGCGCGCAGGCGGTAGCCCGATCGCTTGACTGGTCCAAAGTCAAC	1557
QY	421	TCACCCCTGACATCTCTGTGSCCTCTGCACCTAAATTTAGTCTGTATGACGCCCCCTT	480	Db	1501	GTGAATGTTGCGACTGGCGCGCAGGCGGTAGCCCGATCGCTTGACTGGTCCAAAGTCAAC	1560
Db	421	TCACCCCTGACATCTCTGTGSCCTCTGCACCTAAATTTAGTCTGTATGACGCCCCCTT	480	QY	1558	CTCGAGCGGGCGCCCTCCCGACTGTTGAGCAATATTCAGACATTTCTTTGTGTCTCCCG	1617
QY	481	AATCGGCTCTCGCGCTGCGGAGCGGTACTTAATCTACATTAATGGCCCAAGAGGCTCC	540	Db	1561	CTCGAGCGGGCGCCCTCCCGACTGTTGAGCAATATTCAGACATTTCTTTGTGTCTCCCG	1620
Db	481	AATCGGCTCTCGCGCTGCGGAGCGGTACTTAATCTACATTAATGGCCCAAGAGGCTCC	540	QY	1618	CTTCGTGGCAAGCTCTCTCTTTTGGGAGCGCGGCACAAACAAAGACAGGTTATCCTTATAAT	1677
QY	541	AATATGACAGTACCGGGTGGCGCGCTACTATCGGTTACGGCCCTTAGTGCCTAAT	600	Db	1621	CTTCGTGGCAAGCTCTCTCTTTTGGGAGCGCGGCACAAACAAAGACAGGTTATCCTTATAAT	1680
Db	541	AATATGACAGTACCGGGTGGCGCGCTACTATCGGTTACGGCCCTTAGTGCCTAAT	600	QY	1678	TATAATCTACTGCTAGTGACACAGATCTGATTGAAAAATGCTCGCGGCCATCGGGTCGCC	1737
QY	601	GCAGTTGAGGCTATGCTATATCCATTTCTTTCTGGCCTCAAAACCAACCAACCCCTACA	660	Db	1681	TATAATCTACTGCTAGTGACACAGATCTGATTGAAAAATGCTCGCGGCCATCGGGTCGCC	1740
Db	601	GCAGTTGAGGCTATGCTATATCCATTTCTTTCTGGCCTCAAAACCAACCAACCCCTACA	660	QY	1738	ATTTCACCTATACACACAGGCTTGGGCGCGGTTCGGTTCGCCATTTCTGGCGCGCGGTT	1797
QY	661	TTCTTTGACATGAATTTCCATTTCTCCACTGATGTGAGGATTTCTTTCAACCTGGCATA	720	Db	1741	ATTTCACCTATACACACAGGCTTGGGCGCGGTTCGGTTCGCCATTTCTGGCGCGCGGTT	1800
Db	661	TTCTTTGACATGAATTTCCATTTCTCCACTGATGTGAGGATTTCTTTCAACCTGGCATA	720	QY	1798	TTGGCTCCACGCTCCGCGCTGCTGCTGGAGGATACATTTGATTAATCGGGCGGGCG	1857
QY	721	GCATCTGAATTTGGTCAATCCAGCGAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG	780	Db	1801	TTGGCTCCACGCTCCGCGCTGCTGCTGGAGGATACATTTGATTAATCGGGCGGGCG	1860
Db	721	GCATCTGAATTTGGTCAATCCAGCGAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG	780	QY	1858	CACACATTTGATGACTTCTGCCCTGAATGCGCGCTTTAGGCCCTCAGGGTTCGTCTTC	1917
QY	781	GTGAGACATCTGGTGTGCTGAGGAGGAAGCCACCTCCGCTCTTGTCAATGTTATGCATA	840	Db	1861	CACACATTTGATGACTTCTGCCCTGAATGCGCGCTTTAGGCCCTCAGGGTTCGTCTTC	1920
Db	781	GTGAGACATCTGGTGTGCTGAGGAGGAAGCCACCTCCGCTCTTGTCAATGTTATGCATA	840	QY	1918	CAGTCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAAGTGGGTAAACCTCGGGAGTTG	1977
QY	841	CATGGCTCTCCAGTTAACTCTATACCAATACCCCTTATACGGTGCCCTTTGGTACTG	900	Db	1921	CAGTCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAAGTGGGTAAACCTCGGGAGTTG	1980
Db	841	CATGGCTCTCCAGTTAACTCTATACCAATACCCCTTATACGGTGCCCTTTGGTACTG	900	QY	1978	TAG 1980	
QY	901	GACTTTGCTTAGAGCTTGATTTCCGAATCTCACCACCTGTAAACCAATACACGTTG	960	Db	1981	TAG 1983	
Db	901	GACTTTGCTTAGAGCTTGATTTCCGAATCTCACCACCTGTAAACCAATACACGTTG	960				
QY	961	TCCCGTTACTCCAGCACTGCTCGTCAAC--TCCGCCCGAGGGCGGACGCGGACTGCGGAG	1017				
Db	961	TCCCGTTACTCCAGCACTGCTCGTCAAC--TCCGCCCGAGGGCGGACGCGGACTGCGGAG	1020				
QY	1018	CTGACCAACATGACGACCAACAGGTTCAAGAAATCTCCACTTTACCGGCTTAATGGG	1077				
Db	1021	CTGACCAACATGACGACCAACAGGTTCAAGAAATCTCCACTTTACCGGCTTAATGGG	1080				
QY	1078	GTAGGTGAAGTCCGCGGGATAGCTTAACATTAACCTTTGCTGACAGCTCCTC	1137				
Db	1081	GTAGGTGAAGTCCGCGGGATAGCTTAACATTAACCTTTGCTGACAGCTCCTC	1140				
QY	1138	GGCGGGCTCCCGACAGAAATTAATTTCTGTCGGCTGGCGGCACTGTTTTATTTCCCGCGCG	1197				

RESULT 15
PCT-US95-13703-2
; Sequence 2, Application PC/TUS9513703
; GENERAL INFORMATION:
; APPLICANT: HEPATITIS E VIRUS ANTIGENS AND
; TITLE OF INVENTION: USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA

COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13703
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.41
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2058 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)
INDIVIDUAL ISOLATE: ORF-2 region
PCT-US95-13703--2

Query Match 98.9%; Score 1957.4; DB 5; Length 2058;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1974; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Qy	1	ATGCGCCCTAGCGCTCTTTTCTGTTGTTCTCTTTCTGTTTCTGCTTATGTTGCGCGGCA	60
Db	1	ATGCGCCCTAGCGCTCTTTTCTGTTGTTCTCTTTTCTGCTTATGTTGCGCGGCA	60
Qy	61	CCGACCGGTGACCGGCTCTGCGCGGCTGTTGCGCGGCGGCGGCGGCGGCGGCGGCGG	120
Db	61	CCGACCGGTGACCGGCTCTGCGCGGCTGTTGCGCGGCGGCGGCGGCGGCGGCGGCGG	120
Qy	121	TTCTGGGGTGACCGGCTCTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGG	180
Db	121	TTCTGGGGTGACCGGCTCTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGG	180
Qy	181	CCCTTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGG	240
Db	181	CCCTTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGG	240
Qy	241	CGGCACTTGGCTGCTGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTT	300
Db	241	CGGCACTTGGCTGCTGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTT	300
Qy	301	CGACCTGCGGCTGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCG	360
Db	301	CGACCTGCGGCTGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCG	360
Qy	361	CCGCTCCCGGAGTTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGG	420
Db	361	CCGCTCCCGGAGTTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGG	420
Qy	421	TCACCCCTGACATCTCTGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCT	480
Db	421	TCACCCCTGACATCTCTGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCT	480
Qy	481	AATCCGCTCTGCGGCTGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCT	540
Db	481	AATCCGCTCTGCGGCTGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGGCT	540
Qy	541	AATTATGACAGTACCGGCTGCGGCTGTTGCGCGGCTGTTGCGCGGCTGTTGCGCGG	600

Db	541	AATTATGACAGTACCGGCTGCGGCTGTTGCGGCTGTTGCGGCTGTTGCGGCTGTTGCGG	600
Qy	601	GCAGTTGAGGCTATGCTATATPCCATTTCTTTCTGGCTTCAAAACCAACCAACCAACCA	660
Db	601	GCAGTTGAGGCTATGCTATATPCCATTTCTTTCTGGCTTCAAAACCAACCAACCAACCA	660
Qy	661	TCTGTTGACATGAATTCATTTCTTCCACTGATGTCAGATTTCTTGTTCACCTGCGCAT	720
Db	661	TCTGTTGACATGAATTCATTTCTTCCACTGATGTCAGATTTCTTGTTCACCTGCGCAT	720
Qy	721	GCATCTGAATTTGCTCATCCCAAGGAGGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG	780
Db	721	GCATCTGAATTTGCTCATCCCAAGGAGGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG	780
Qy	781	GTTGAGACATCTGTTGTTGCTGAGGAGGAGCACTCCCGTCTTGTCAITGATGTCATA	840
Db	781	GTTGAGACATCTGTTGTTGCTGAGGAGGAGCACTCCCGTCTTGTCAITGATGTCATA	840
Qy	841	CATGCTCTCCAGTTAACTCCTATACCAATACCCCTTATACGGTGCCCTTGGCTTACTG	900
Db	841	CATGCTCTCCAGTTAACTCCTATACCAATACCCCTTATACGGTGCCCTTGGCTTACTG	900
Qy	901	GACTTTGCTTTAGAGCTTTGAGTTTCGCAATCTCACCACTGTAAACCAATACACGCTG	960
Db	901	GACTTTGCTTTAGAGCTTTGAGTTTCGCAATCTCACCACTGTAAACCAATACACGCTG	960
Qy	961	TCCGTTTACTCCAGACTGCTGCTGAC---TCGCGCGGAGGCGGCGGAGCTGCGGAG	1017
Db	961	TCCGTTTACTCCAGACTGCTGCTGAC---TCGCGCGGAGGCGGCGGAGCTGCGGAG	1020
Qy	1018	CTGACCAACTGCGAGCCAGGTTTCATGAAGATCTCCACTTTACCGGCTTAAATGGG	1077
Db	1021	CTGACCAACTGCGAGCCAGGTTTCATGAAGATCTCCACTTTACCGGCTTAAATGGG	1080
Qy	1078	GTAGGTGAAGTCCGCGCGGATAGCTCTAACTTTACTTAACTTGTGTGACACGCTCCTC	1137
Db	1081	GTAGGTGAAGTCCGCGCGGATAGCTCTAACTTTACTTAACTTGTGTGACACGCTCCTC	1140
Qy	1138	GGCGGCTCCGCAAGAAATTAATTTGCTGCGCTGGCGGCAACTGTTTTATTTCGCGCGC	1197
Db	1141	GGCGGCTCCGCAAGAAATTAATTTGCTGCGCTGGCGGCAACTGTTTTATTTCGCGCGC	1200
Qy	1198	GTTGCTCAGCAATGCGAGCCAGCTGAGCTTATACATCAGTGGAGAAATGCTCAG	1257
Db	1201	GTTGCTCAGCAATGCGAGCCAGCTGAGCTTATACATCAGTGGAGAAATGCTCAG	1260
Qy	1258	CAGGATAAGGCTGTTGCTATCCCGCAGCATATCGATCTTGTGTGATTCGCGTGTGCTAT	1317
Db	1261	CAGGATAAGGCTGTTGCTATCCCGCAGCATATCGATCTTGTGTGATTCGCGTGTGCTAT	1320
Qy	1318	CAGGATATGACAAACGAGATGAGCGATGCGGCCACCGCTGCGCTGCGCATCTCGG	1377
Db	1321	CAGGATATGACAAACGAGATGAGCGATGCGGCCACCGCTGCGCTGCGCATCTCGG	1380
Qy	1378	CTTTTCTGTTCTCCGAGCAATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1437
Db	1381	CTTTTCTGTTCTCCGAGCAATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1440
Qy	1438	GACAGTCCACTTACCGGCTGCTCAACTGCGCGGCTTATATCTCGGACAGCGTACTTTG	1497
Db	1441	GACAGTCCACTTACCGGCTGCTCAACTGCGCGGCTTATATCTCGGACAGCGTACTTTG	1500
Qy	1498	GTGAATGTTGCGACTGCGCGGCGGCGTATGCGCGATGCTGCTGCTGCTGCTGCTGCTG	1557
Db	1501	GTGAATGTTGCGACTGCGCGGCGGCGTATGCGCGATGCTGCTGCTGCTGCTGCTGCTG	1560
Qy	1558	CTGACGGGCGGCGGCTCCCGACTGTTGAGCAATATTCAGACATTTCTTTGCTGCTCCC	1617
Db	1561	CTGACGGGCGGCGGCTCCCGACTGTTGAGCAATATTCAGACATTTCTTTGCTGCTCCC	1620
Qy	1618	CTTCGTGGCAAGCTCTCTTTTGGGAGCGGCGGCAACAAAGCAGGTTATCTTATAAT	1677
Db	1621	CTTCGTGGCAAGCTCTCTTTTGGGAGCGGCGGCAACAAAGCAGGTTATCTTATAAT	1680

Qy 1678 TATATACTACTGCTAGTCACACAGATTCTGATTGAATGCTGCGGCCATCGGTCGCC 1737
Db |||||
1681 TATAATACTACTGCTAGTCACACAGATTCTGATTGAATAATGCTGCGGCCATCGGTCGCC 1740
Qy 1738 ATTTCAACCTATACACACAGGCTTGGGGCCGGTCCGGTCGCCATTTCTGCGGCCCGGGTT 1797
Db |||||
1741 ATTTCAACCTATACACACAGGCTTGGGGCCGGTCCGGTCGCCATTTCTGCGGCCCGGGTT 1800
Qy 1798 TTGGTCCACGCTCCGCGCCCTGGCTCTGCTGGAGGATACCTTTTGATTATCCGGGGCGGGCG 1857
Db |||||
1801 TTGGTCCACGCTCCGCGCCCTGGCTCTGCTGGAGGATACCTTTTGATTATCCGGGGCGGGCG 1860
Qy 1858 CACACATTGTGATGACTTCTGCGCTGAATGCGCGCTTTAGGCTCCAGGGTTGTGCTTTC 1917
Db |||||
1861 CACACATTGTGATGACTTCTGCGCTGAATGCGCGCTTTAGGCTCCAGGGTTGTGCTTTC 1920
Qy 1918 CAGTCAACTGTCGCTGAGCTCCAGCGCCTTAAAGTTAAGTGGGTAAACTCGGGAGTTG 1977
Db |||||
1921 CAGTCAACTGTCGCTGAGCTCCAGCGCCTTAAAGTTAAGTGGGTAAACTCGGGAGTTG 1980
Qy 1978 TAG 1980
Db |||
1981 TAG 1983

Search completed: August 21, 2004, 18:11:08
Job time : 167.394 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2004, 15:15:13 ; Search time 885,329 Seconds
(without alignments)
10985.769 Million cell updates/sec

Title: US-09-851-410a-10_COPY_5117_7096

Perfect score: 1980

Sequence: 1 ATGGCCCTAGGCTCTTTT.....GTAAACTCGGAGTTGAG 1980

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 322839 seqs, 245606551 residues

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1980	100.0	2055	15	US-10-165-868-4
2	1980	100.0	2100	15	US-10-165-868-2
3	1980	100.0	7171	10	US-09-851-410-10
4	1980	100.0	7180	17	US-10-239-090A-49
5	1957.4	98.9	2058	9	US-09-769-066-2
6	1621.4	81.9	1647	9	US-09-769-066-4
7	1371.8	69.3	1990	13	US-10-381-770-5
8	1368.6	69.1	2049	9	US-09-769-066-1
9	1368.6	69.1	2049	15	US-10-165-868-3
10	1368.6	69.1	2094	15	US-10-165-868-1
11	1368.6	69.1	7195	10	US-09-851-410-6
12	1361.4	68.8	1989	13	US-10-381-770-6
13	1319.6	66.6	7256	17	US-10-239-090A-44
14	1315.2	66.4	7277	10	US-09-468-147-164

15	1315.2	66.4	7277	10	US-09-468-147-165	Sequence 165, App
16	1315.2	66.4	7277	13	US-10-319-745-164	Sequence 164, App
17	1315.2	66.4	7277	13	US-10-319-745-165	Sequence 165, App
18	1314	66.4	7228	17	US-10-239-090A-47	Sequence 47, Appl
19	1313.2	66.3	7233	17	US-10-239-090A-45	Sequence 45, Appl
20	1308	66.1	1308	15	US-10-165-868-10	Sequence 10, Appl
21	1301.2	65.7	7229	17	US-10-239-090A-46	Sequence 46, Appl
22	1301.2	65.7	7234	17	US-10-239-090A-11	Sequence 11, Appl
23	1278	64.5	7230	17	US-10-239-090A-48	Sequence 48, Appl
24	1274	64.3	7202	10	US-09-468-147-89	Sequence 89, Appl
25	1274	64.3	7202	10	US-09-468-147-90	Sequence 90, Appl
26	1274	64.3	7202	13	US-10-319-745-89	Sequence 89, Appl
27	1274	64.3	7202	13	US-10-319-745-90	Sequence 90, Appl
28	1090.2	55.1	1647	9	US-09-769-066-3	Sequence 3, Appl
29	984	49.7	984	9	US-09-769-066-6	Sequence 6, Appl
30	981	49.5	981	15	US-10-165-868-8	Sequence 8, Appl
31	880.6	44.5	1311	15	US-10-165-868-9	Sequence 9, Appl
32	673.6	34.0	984	9	US-09-769-066-5	Sequence 5, Appl
33	673.6	34.0	984	15	US-10-165-868-7	Sequence 7, Appl
34	660.8	33.4	1020	10	US-09-468-147-129	Sequence 129, App
35	660.8	33.4	1020	13	US-10-319-745-129	Sequence 129, App
36	640.6	32.4	1389	10	US-09-468-147-193	Sequence 193, App
37	640.6	32.4	1389	10	US-09-468-147-196	Sequence 196, App
38	640.6	32.4	1389	13	US-10-319-745-193	Sequence 193, App
39	640.6	32.4	1389	13	US-10-319-745-196	Sequence 196, App
40	638.2	32.2	1026	10	US-09-468-147-192	Sequence 192, App
41	638.2	32.2	1026	10	US-09-468-147-195	Sequence 195, App
42	638.2	32.2	1026	13	US-10-319-745-192	Sequence 192, App
43	638.2	32.2	1026	13	US-10-319-745-195	Sequence 195, App
44	558.2	28.2	899	10	US-09-468-147-24	Sequence 24, Appl
45	558.2	28.2	899	13	US-10-319-745-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-10-165-868-4
; Sequence 4, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Twu, Jr., Shin
; Bradley, Daniel W.
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10165.868
; FILING DATE: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870.985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921

QY 901 GACTTTGCTTAGAGCTTGGATTTTCGCAATCTCACCACTGTAAACCAATACAGTGTG 960
Db 946 GACTTTGCTTAGAGCTTGGATTTTCGCAATCTCACCACTGTAAACCAATACAGTGTG 1005
QY 961 TCCCGTTACTTCAGCACTGCTCGTCACTCCGCGCGAGGGCGCAGCGGACTGCGGAGTGT 1020
Db 1006 TCCCGTTACTTCAGCACTGCTCGTCACTCCGCGCGAGGGCGCAGCGGACTGCGGAGTGT 1065
QY 1021 ACCCAACTGACGACCAACAGGTTATGAAGAATCTCCATTTTACCGGCTTAATGGGTA 1080
Db 1066 ACCCAACTGACGACCAACAGGTTATGAAGAATCTCCATTTTACCGGCTTAATGGGTA 1125
QY 1081 GGTGAAGTCGGCGCGGATAGCTCTAACATTACTTAACCTTGTGACACAGCTCTCTCGC 1140
Db 1126 GGTGAAGTCGGCGCGGATAGCTCTAACATTACTTAACCTTGTGACACAGCTCTCTCGC 1185
QY 1141 GGGCTCCCGACAGAATAATTTGCTCGGCTGGCGGCAACTGTTTATTTCCGCGCGGTT 1200
Db 1186 GGGCTCCCGACAGAATAATTTGCTCGGCTGGCGGCAACTGTTTATTTCCGCGCGGTT 1245
QY 1201 GTCTCAGCAATGGCGAGCAACCGTGAAGCTCTATACATCATGTGAGAAATCTCAGCAG 1260
Db 1246 GTCTCAGCAATGGCGAGCAACCGTGAAGCTCTATACATCATGTGAGAAATCTCAGCAG 1305
QY 1261 GATAGGGTGTGCTATATCCCGACGATATCGATCTTGGTGATTTCGGGTGTGGTCAATTCAG 1320
Db 1306 GATAGGGTGTGCTATATCCCGACGATATCGATCTTGGTGATTTCGGGTGTGGTCAATTCAG 1365
QY 1321 GATTATGACAAACAGCATGAGCAGGATCGGCCACCCCGTCCGCTCGCCCATCTCGGCT 1380
Db 1366 GATTATGACAAACAGCATGAGCAGGATCGGCCACCCCGTCCGCTCGCCCATCTCGGCT 1425
QY 1381 TTTTCTGTTCTCGAGCAAAATGATGATCTTGGTGTCTCTCACTCAGCGGATGAC 1440
Db 1426 TTTTCTGTTCTCGAGCAAAATGATGATCTTGGTGTCTCTCACTCAGCGGATGAC 1485
QY 1441 CAGTCCACTTACGGTCTGCACTGCGCGGTTTATATCTCGGACAGCGTGAATTTGGTG 1500
Db 1486 CAGTCCACTTACGGTCTGCACTGCGCGGTTTATATCTCGGACAGCGTGAATTTGGTG 1545
QY 1501 AATGTTGCACTGGCGCGAGGCGGTAGCCCGATCGCTTGACTGTGTCAAAGTCACGCTC 1560
Db 1546 AATGTTGCACTGGCGCGAGGCGGTAGCCCGATCGCTTGACTGTGTCAAAGTCACGCTC 1605
QY 1561 GACGGCGGCGCTCCGACTGTGAGCAATATTCAGACATTTCTTGTGCTCCGCTT 1620
Db 1606 GACGGCGGCGCTCCGACTGTGAGCAATATTCAGACATTTCTTGTGCTCCGCTT 1665
QY 1621 CGTGGCAAGCTCTCTTTTGGAGCGCGGACAAACAAAGCAGGTTATCTTATAATTAT 1680
Db 1666 CGTGGCAAGCTCTCTTTTGGAGCGCGGACAAACAAAGCAGGTTATCTTATAATTAT 1725
QY 1681 AATGTTGCACTGGCGCGAGGCGGTAGCCCGATCGCTTGACTGTGTCAAAGTCACGCTC 1740
Db 1726 AATGTTGCACTGGCGCGAGGCGGTAGCCCGATCGCTTGACTGTGTCAAAGTCACGCTC 1785
QY 1741 TCAACTATACACAGGCTTGGCGCGGTTCGGTCCGCTTCTCGGCGCGGTTTG 1800
Db 1786 TCAACTATACACAGGCTTGGCGCGGTTCGGTCCGCTTCTCGGCGCGGTTTG 1845
QY 1801 GCTCCACGCTCCGCGCTGCTGCTGGAGGATACATTTGATTATCCGGCGCGGCGCAC 1860
Db 1846 GCTCCACGCTCCGCGCTGCTGCTGGAGGATACATTTGATTATCCGGCGCGGCGCAC 1905
QY 1861 ACATTTGATGACTTTCGCTTGAAGTCGCGGCTTTAGGCTTCCAGGTTGTGCTTCCAG 1920
Db 1906 ACATTTGATGACTTTCGCTTGAAGTCGCGGCTTTAGGCTTCCAGGTTGTGCTTCCAG 1965
QY 1921 TCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGTGAAGAACTCGGAGTGTAG 1980
Db 1966 TCAACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGTGAAGAACTCGGAGTGTAG 2025

RESULT 3
US-09-851-410-10
; Sequence 10, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yarbough, Patrice O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; No. US20030124510A1-A/No. US20030124510A1-B Hepatitis Viral
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,410
; FILING DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0980
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1717 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: Composite Mexico strain
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-851-410-10

Query Match 100.0%; Score 1980; DB 10; Length 7171;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGCGCCCTAGGCTCTTTTGTGCTTGTCTCTTGTCTGCTATGTTGCCGCGCCA 60
Db 5117 ATCGCGCCCTAGGCTCTTTTGTGCTTGTCTCTTGTCTGCTATGTTGCCGCGCCA 5176
QY 61 CCGACCGGTGAGCGGCTCTGCGCGCGCTGTTGGGCGCGCAGCGCGGTACCGGCGGTGGT 120

Db 5177 CCGACCGGTACCGCTCTGGCCGCGTGTGGCGGGGACGGGGTACCGCGGTGGT 5236
QY 121 TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGGAATCCCGCTATATTCATCCAAACAC 180
Db 5237 TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGGAATCCCGCTATATTCATCCAAACAC 5296
QY 181 CCCTTTGGCCCGACAGCTTGGCGCTGCGTCCGGGTCTGGACCTCGCTTCGCGCAACACGAC 240
Db 5297 CCCTTTGGCCCGACAGCTTGGCGCTGCGTCCGGGTCTGGACCTCGCTTCGCGCAACACGAC 5356
QY 241 CGGCCACTTGGCTCCACTTGGCGGAGATCAGGCCCGAGGCCCTTCGCTCCGCTCCGCTGCG 300
Db 5357 CGGCCACTTGGCTCCACTTGGCGGAGATCAGGCCCGAGGCCCTTCGCTCCGCTCCGCTGCG 5416
QY 301 CGACCTGCCACAGCGGGGCTGCGCGCTGACGGCTGAGCGCTGAGCGCTGCCCATGACACCTCA 360
Db 5417 CGACCTGCCACAGCGGGGCTGCGCGCTGACGGCTGAGCGCTGAGCGCTGCCCATGACACCTCA 5476
QY 361 CCCGTCGGGACGTTGATTTCTCGCGGTGCAATTTCTAGCGGCCAGTATAATTTCTCTACT 420
Db 5477 CCCGTCGGGACGTTGATTTCTCGCGGTGCAATTTCTAGCGGCCAGTATAATTTCTCTACT 5536
QY 421 TCACCCCTGACATCCTCTGTGGCCCTCTGGCACTAATTTAGTCTGTATGCAAGCCCGCCT 480
Db 5537 TCACCCCTGACATCCTCTGTGGCCCTCTGGCACTAATTTAGTCTGTATGCAAGCCCGCCT 5596
QY 481 AATCCGCTCTGCGCTGCAGGACGGTACTTAATCTACATATATGCGCCACAGAGGCTCC 540
Db 5597 AATCCGCTCTGCGCTGCAGGACGGTACTTAATCTACATATATGCGCCACAGAGGCTCC 5656
QY 541 AATATGCAAGTACCGGGTTGCGCGCTACTATCGTTTACCGGCCCTAGTGCCTAAT 600
Db 5657 AATATGCAAGTACCGGGTTGCGCGCTACTATCGTTTACCGGCCCTAGTGCCTAAT 5716
QY 601 GCAAGTGAGGCTAGTATATATCAATTTCTTTCTGGCCCTCAACCAACCAACCCCTACA 660
Db 5717 GCAAGTGAGGCTAGTATATATCAATTTCTTTCTGGCCCTCAACCAACCAACCCCTACA 5776
QY 661 TCTGTTGACATGAATTCATTAATCTTCACTGATGTGAGGATCTTGTTCACCTGGGATA 720
Db 5777 TCTGTTGACATGAATTCATTAATCTTCACTGATGTGAGGATCTTGTTCACCTGGGATA 5836
QY 721 GCATCTGAATTTGGTCAATCCCAAGGAGCGCTTCACTACCGCAATCAAGGTTGGCGTCC 780
Db 5837 GCATCTGAATTTGGTCAATCCCAAGGAGCGCTTCACTACCGCAATCAAGGTTGGCGTCC 5896
QY 781 GTTGAGACATCTGGTGTGTGAGGAGGAGCCACCTCCGCTCTTGTCAATGTTATGCAATA 840
Db 5897 GTTGAGACATCTGGTGTGTGAGGAGGAGCCACCTCCGCTCTTGTCAATGTTATGCAATA 5956
QY 841 CATGGCTCTCCAGTTAACTCTATACCAATACCGCTTATACCGTGGCCCTTGGCTTACTG 900
Db 5957 CATGGCTCTCCAGTTAACTCTATACCAATACCGCTTATACCGTGGCCCTTGGCTTACTG 6016
QY 901 GACTTTGCCCTTAGAGCTTAGTTGCGCAATCTACACCACTGTAAACCAATACACGTTG 960
Db 6017 GACTTTGCCCTTAGAGCTTAGTTGCGCAATCTACCACTGTAAACCAATACACGTTG 6076
QY 961 TCCCGTTACTCCAGCACTGCTGCTACTCCGCGCGAGGGGCGGAGGACTCGGAGCTG 1020
Db 6077 TCCCGTTACTCCAGCACTGCTGCTACTCCGCGCGAGGGGCGGAGGACTCGGAGCTG 6136
QY 1021 ACCAATCTGACGACAGGTTTCAATGAAGATCTCACTTACCGGCTTAAATGGGGTA 1080
Db 6137 ACCAATCTGACGACAGGTTTCAATGAAGATCTCACTTACCGGCTTAAATGGGGTA 6196
QY 1081 GGTGAAGTCCGCGCGGAGTGTCTTAACATTAATTAACCTTGTGACAGCTCTCTCGG 1140
Db 6197 GGTGAAGTCCGCGCGGAGTGTCTTAACATTAATTAACCTTGTGACAGCTCTCTCGG 6256
QY 1141 GGGCTCCCGACAGAAATTAATTTCTGCGGCTGGCGGCAACTGTTTATTTCCCGCCCGGTT 1200

Db 6257 GGGCTCCCGACAGAAATTAATTTCTGTCGCTGCGGGGCAACTGTTTATTTCCCGCCCGGTT 6316
QY 1201 GTCTCAGCCAATGCGGAGCCCAACCGTGAAGCTCTATATCATCAGTGGAGAATGCTCAGCAG 1260
Db 6317 GTCTCAGCCAATGCGGAGCCCAACCGTGAAGCTCTATATCATCAGTGGAGAATGCTCAGCAG 6376
QY 1261 GATAAGGGTGTGTCTATCCCGCCAGATATCGATCTTGCTGATTCGCTGTCGCTCATTCAG 1320
Db 6377 GATAAGGGTGTGTCTATCCCGCCAGATATCGATCTTGCTGATTCGCTGTCGCTCATTCAG 6436
QY 1321 GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCGCTGCGCTGCGCCACTCTCGGCT 1380
Db 6437 GATTATGACAAACAGCATGAGCAGGATCGGCCCAACCGCTGCGCTGCGCCACTCTCGGCT 6496
QY 1381 TTTTCTGTTCTCCGAGCAATGATGATGTTGGCTGTCCCTCACTGAGCCGAGTATGAC 1440
Db 6497 TTTTCTGTTCTCCGAGCAATGATGATGTTGGCTGTCCCTCACTGAGCCGAGTATGAC 6556
QY 1441 CAGTCCACTTACGGGTCTGCAACTGGCCCGGTTTATATCTCGGACAGCTGACTTTGGTG 1500
Db 6557 CAGTCCACTTACGGGTCTGCAACTGGCCCGGTTTATATCTCGGACAGCTGACTTTGGTG 6616
QY 1501 AATGTTCCGACTGCGCGCAGCGCGCTAGCCGATCGCTTGAATGCAATTTCTTTGCTCCCGCT 1560
Db 6617 AATGTTCCGACTGCGCGCAGCGCGCTAGCCGATCGCTTGAATGCAATTTCTTTGCTCCCGCT 6676
QY 1561 GACGGGGGGCCCTCCCGACTGTTGAGCAATATTTCAAAGACATTTCTTTGCTCCCGCT 1620
Db 6677 GACGGGGGGCCCTCCCGACTGTTGAGCAATATTTCAAAGACATTTCTTTGCTCCCGCT 6736
QY 1621 CGTGGCAAGCTCTCTTTTGGGAGCGCGCACAAAGACAGGTTATCCTTTATAATAT 1680
Db 6737 CGTGGCAAGCTCTCTTTTGGGAGCGCGCACAAAGACAGGTTATCCTTTATAATAT 6796
QY 1681 AATACTACTGCTAGTAGCAGATTTCTGATTAATAATGCTGCGGCCCATTCGGGTGCCCAT 1740
Db 6797 AATACTACTGCTAGTAGCAGATTTCTGATTAATAATGCTGCGGCCCATTCGGGTGCCCAT 6856
QY 1741 TCAACCTATACCAAGGCTTGGGGCGGCTCGGTCGCTGCGCATTTCTGCGGCGCGGTTTG 1800
Db 6857 TCAACCTATACCAAGGCTTGGGGCGGCTCGGTCGCTGCGCATTTCTGCGGCGCGGTTTG 6916
QY 1801 GCTCCAGCTCCGCGCTCGCTCTGCTGAGGATATTTTGAATTAATCCGGGCGGCGCAC 1860
Db 6917 GCTCCAGCTCCGCGCTCGCTCTGCTGAGGATATTTTGAATTAATCCGGGCGGCGCAC 6976
QY 1861 ACATTTGATGACTTCTGCCCTGAATGCCGCTTTAGGCCCTCCAGGTTGTGCTTCCAG 1920
Db 6977 ACATTTGATGACTTCTGCCCTGAATGCCGCTTTAGGCCCTCCAGGTTGTGCTTCCAG 7036
QY 1921 TCACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTGTAG 1980
Db 7037 TCACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTGTAG 7096

RESULT 4

US-10-239-090A-49
; Sequence 49, Application US/10239090A
; Publication No. US20040101820A1
; GENERAL INFORMATION:
; APPLICANT: KABUSHIKI KAISHA TOSHIBA
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus col
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same, i
; FILE REFERENCE: 02S0741P
; CURRENT APPLICATION NUMBER: US/10/239,090A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 49
; LENGTH: 7180
; TYPE: DNA

Db 1441 GACGAGTCCACTTACGGGTGTCACCTGGCCGGGTTTATATCTCGGACAGCGTGACTTTG 1500
 QY 1498 GTGAATGTGGACTGGCGGCGAGCCGTAGCCGATCGCTTGACTGCTGCACAAAGTCAACC 1557
 Db 1501 GTGAATGTGGACTGGCGGCGAGCCGTAGCCGATCGCTTGACTGCTGCACAAAGTCAACC 1560
 QY 1558 CTGACGGCGGCGCCCTCCGACTCTTCAGCAATATTCAGACATTTCTTTGTGCTCCCC 1617
 Db 1561 CTGACGGCGGCGCCCTCCGACTCTTCAGCAATATTCAGACATTTCTTTGTGCTCCCC 1620
 QY 1618 CTTCTGGCAAGCTCTCTTTTGGAGGCGCGCAACAACAAAGCAGGTTATCTTATAT 1677
 Db 1621 CTTCTGGCAAGCTCTCTTTTGGAGGCGCGCAACAACAAAGCAGGTTATCTTATAT 1680
 QY 1678 TATAATCTACTGTAGTACAGATTTCTGATTGAATGCTGCGGCGCATCGGTCGCC 1737
 Db 1681 TATAATCTACTGTAGTACAGATTTCTGATTGAATGCTGCGGCGCATCGGTCGCC 1740
 QY 1738 ATTTCAACCTATACACAGGCTTGGGCGCGGTCCCGTCCCATTTCTGCGGCGCGGTT 1797
 Db 1741 ATTTCAACCTATACACAGGCTTGGGCGCGGTCCCGTCCCATTTCTGCGGCGCGGTT 1800
 QY 1798 TTGGCTCCACGCTCCGCGCTCTGCTGAGGATACCTTTTGAATATCCGGGCGCGG 1857
 Db 1801 TTGGCTCCACGCTCCGCGCTCTGCTGAGGATACCTTTTGAATATCCGGGCGCGG 1860
 QY 1858 CACATATTTGAGCTTCTGCGCTGATGCGGCTTTAGGCTCCAGGCTTGTGCTTTC 1917
 Db 1861 CACATATTTGAGCTTCTGCGCTGATGCGGCTTTAGGCTCCAGGCTTGTGCTTTC 1920
 QY 1918 CAGTCAACTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGTAAAGTCTGGGAGTTG 1977
 Db 1921 CAGTCAACTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGTAAAGTCTGGGAGTTG 1980
 QY 1978 TAG 1980
 Db 1981 TAG 1983

RESULT 6

US-09-769-066-4
 ; Sequence 4, Application US/09769066
 ; Patent No. US20020107360A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fuerst, Thomas R.
 ; McAtee, C. Patrick
 ; Yarbrough, Patrice O.
 ; Zhang, Yifan
 ; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: 350 Cambridge Ave., Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/769,066
 ; FILING DATE: 24-Jan-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/542,634
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fabian, Gary R.
 ; REGISTRATION NUMBER: 33,875
 ; REFERENCE/DOCKET NUMBER: 4600-0293.30

TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 324-0880
 ; TELEFAX: (415) 324-0960
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1647 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: Hepatitis E virus (Mexico strain)
 ; r62kDa, FIGURE 2
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 US-09-769-066-4
 Query Match 81.9%; Score 1621.4; DB 9; Length 1647;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1638; Conservative 0; Mismatches 6; Indels 3; Gaps 1;
 QY 334 GCTGTGGGCGCTGCCCATGACACCTCACCCGTCGCGAGCGTTGATTCGCGGTGCAATT 393
 Db 1 GCTGTGGGCGCTGCCCATGACACCTCACCCGTCGCGAGCGTTGATTCGCGGTGCAATT 60
 QY 394 CTACGCGCGCAGTATAATTTGTCTACTTCAACCCCTGACATCTCTGTGCGCTCTGCACT 453
 Db 61 CTACGCGCGCAGTATAATTTGTCTACTTCAACCCCTGACATCTCTGTGCGCTCTGCACT 120
 QY 454 RAATTTAGTCTGTATGACGCGCCCTTAATCGCGCTCTGCGCTGCGAGCGGTACTAAT 513
 Db 121 AATTTAGTCTGTATGACGCGCCCTTAATCGCGCTCTGCGCTGCGAGCGGTACTAAT 180
 QY 514 ACTCACATTATGGCCACAGAGCGCTCCCAATTATGACAGTACCGGTTGCGCGGTACT 573
 Db 181 ACTCACATTATGGCCACAGAGCGCTCCCAATTATGACAGTACCGGTTGCGCGGTACT 240
 QY 574 ATCCGTTACCGGCGCTTAGTCTTAATGAGTTGGAGCTATGCTATATCCATTTCTTTC 633
 Db 241 ATCCGTTACCGGCGCTTAGTCTTAATGAGTTGGAGCTATGCTATATCCATTTCTTTC 300
 QY 634 TGGCGCTCAACACACCAACCCCTACATCTGTTGACATGAATTCATTAATTCCTCACTGAT 693
 Db 301 TGGCGCTCAACACACCAACCCCTACATCTGTTGACATGAATTCATTAATTCCTCACTGAT 360
 QY 694 GTCAGGATTTCTTTTCAACCTGTCATAGCATATGAAATGCTATCCCAAGCGAGCGCTT 753
 Db 361 GTCAGGATTTCTTTTCAACCTGTCATAGCATATGAAATGCTATCCCAAGCGAGCGCTT 420
 QY 754 CACTACCGCAATCAAGTTGGCGCTCGGTTGAGACATCTGGTGTGCTGAGGAGGAGGCC 813
 Db 421 CACTACCGCAATCAAGTTGGCGCTCGGTTGAGACATCTGGTGTGCTGAGGAGGAGGCC 480
 QY 814 ACCTCGGCTTGTGTCATGTTATGACATATGCTCTCCAGTTAACTTAACTTAACTAATACC 873
 Db 481 ACCTCGGCTTGTGTCATGTTATGACATATGCTCTCCAGTTAACTTAACTTAACTAATACC 540
 QY 874 CTTTATACCGGTGCGCTTGTGCTTACTGGACTTTTGCCTTAGAGCTTGGTTTTCGAATCTC 933
 Db 541 CTTTATACCGGTGCGCTTGTGCTTACTGGACTTTTGCCTTAGAGCTTGGTTTTCGAATCTC 600
 QY 934 ACCACCTGTAACACCAATACACGCTGCTCCGTTACTCCAGCACTGCTCGTAC---TCC 990
 Db 601 ACCACCTGTAACACCAATACACGCTGCTCCGTTACTCCAGCACTGCTCGTAC---TCC 660
 QY 991 GCCCGAGGGGCGAGCGGACTGCGGAGCTGACCAAACTGACAGCCACCGAGTTTCATGAAA 1050
 Db 661 CGCCGAGGGGCGAGCGGACTGCGGAGCTGACCAAACTGACAGCCACCGAGTTTCATGAAA 720
 QY 1051 GATCTCCACTTTTACCGGCTTAAATGGGGTAGGTGAAGTCGGCGCGGGAATAGCTCTAACA 1110
 Db 721 GATCTCCACTTTTACCGGCTTAAATGGGGTAGGTGAAGTCGGCGCGGGAATAGCTCTAACA 780
 QY 1111 TTACTTAACCTTGTGACACGCTCTCCGCGGGCTCCCGACAGAAATTAATTTCTCGGCT 1170

```
Db 781 TTACTTAACCTTGCTGACAGCTCTCTCGCGGCTCCCGACAGAAATAATTTCGTCGGCT 840
QY 1171 GCGGGCAACTGTTTTATTCCCGCCGGTGTCTCAGCCAAATGGCGAGCCAAACCTGAG 1230
Db 841 GCGGGCAACTGTTTTATTCCCGCCGGTGTCTCAGCCAAATGGCGAGCCAAACCTGAG 900
QY 1231 CTCTATACATCAGTCGAGAAATCTCAGCAGGATAAGGTGTGCTATCCCCCAGCATATC 1290
Db 901 CTCTATACATCAGTCGAGAAATCTCAGCAGGATAAGGTGTGCTATCCCCCAGCATATC 960
QY 1291 GATCTTGATGATTCGCTGTGTGTCATTCAGGATATGACAAACAGATGAGCAGATCGG 1350
Db 961 GATCTTGATGATTCGCTGTGTGTCATTCAGGATATGACAAACAGATGAGCAGATCGG 1020
QY 1351 CCCACCCGTCGCTCGCGCATCTCGGCCATCTCGGCCATCTCGGCCATCTCGGCCATCT 1410
Db 1021 CCCACCCGTCGCTCGCGCATCTCGGCCATCTCGGCCATCTCGGCCATCTCGGCCATCT 1080
QY 1411 TGGCTGTCCCTCACTGACGCGAGATGACACAGTCCACTTACGGTCTGTCAACTGGCCCG 1470
Db 1081 TGGCTGTCCCTCACTGACGCGAGATGACACAGTCCACTTACGGTCTGTCAACTGGCCCG 1140
QY 1471 GTTTATATCTCGACAGCGTGACTTTGGTGAATGTGCGACTGCGCGAGGCCCTAGCC 1530
Db 1141 GTTTATATCTCGACAGCGTGACTTTGGTGAATGTGCGACTGCGCGAGGCCCTAGCC 1200
QY 1531 CGATCGCTTGACTGCTCCAAAGTCACTTCGACGCGCGGCCCTCCCGACTGTTGAGCAA 1590
Db 1201 CGATCGCTTGACTGCTCCAAAGTCACTTCGACGCGCGGCCCTCCCGACTGTTGAGCAA 1260
QY 1591 TATTCCAAAGACATTTTGTGTCCTCCCTTCGTGGCAAGCTCTCTCTTTGGAGGCCCGC 1650
Db 1261 TATTCCAAAGACATTTTGTGTCCTCCCTTCGTGGCAAGCTCTCTCTTTGGAGGCCCGC 1320
QY 1651 ACAACAAAGCAGGTATCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1710
Db 1321 ACAACAAAGCAGGTATCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
QY 1711 GAAATGCTGCGGCGCATCGGTCGCCATTTCACTTATACATATACCAAGCTTTGGGCGCG 1770
Db 1381 GAAATGCTGCGGCGCATCGGTCGCCATTTCACTTATACATATACCAAGCTTTGGGCGCG 1440
QY 1771 CCGTCGCAATTTCTGCGGCGCGGTTTTGGCTCCACGCTCGGCCCTGGCTCTGCTGGAG 1830
Db 1441 CCGTCGCAATTTCTGCGGCGCGGTTTTGGCTCCACGCTCGGCCCTGGCTCTGCTGGAG 1500
QY 1831 GATACCTTTGATTAATCGGGCGGGCGGACACATTTGATGACTTCTGCCCTGAATGCGCG 1890
Db 1501 GATACCTTTGATTAATCGGGCGGGCGGACACATTTGATGACTTCTGCCCTGAATGCGCG 1560
QY 1891 GCTTTAGGCTCCAGGTTGTCTTCCAGTCACTGTCGCTGAGCTCCAGGCTTAA 1950
Db 1561 GCTTTAGGCTCCAGGTTGTCTTCCAGTCACTGTCGCTGAGCTCCAGGCTTAA 1620
QY 1951 GTTAAGTGGTAAACTCGGGAGTTG 1977
Db 1621 GTTAAGTGGTAAACTCGGGAGTTG 1647
```

```
RESULT 7
US-10-381-770-5
; Sequence 5, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS
; TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
; FILE REFERENCE: IEC010037PCT
; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 5
; LENGTH: 1990
; TYPE: DNA
; ORGANISM: hepatitis E virus
US-10-381-770-5
```

```
Query Match 69.3%; Score 1371.8; DB 13; Length 1990;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 1608; Conservative 0; Mismatches 372; Indels 3; Gaps 1;
```

```
QY 1 ATGCGCCCTAGGCTCTTTTGTGTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
Db 1 ATGCGCCCTCGGCTATTTTGTGTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
QY 61 CCGACCGGTGACGCTCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CCGCGCGGTGACGCTCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 TTCTGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATATTCATCAACCCAC 180
Db 121 TTCTGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCCAC 180
QY 181 CCTTTGCCCGCAGAGTTGCGGCTCGGCTCGGCTCTGGACCTTCGCTTCGCCAAACGAGCC 240
Db 181 CCTTTGCCCGCAGAGTTGCGGCTCGGCTCGGCTCTGGACCTTCGCTTCGCCAAACGAGCC 240
QY 241 CGGCACTTGGCTCCACTTGGCGGAGATCAGGCGCAGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 CGACACTCGGCTCGGCTTGGCGGAGATCAGGCGCAGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 CGACTGCGCACAGCGCGGCTGCGCGCTGACGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 360
Db 301 AGACTGCGCACAGCGCGGCTGCGCGCTGACGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 360
QY 361 CCGCTCCGCGAGTTGATTTCTCGGCTGCAATTTCTACGCGCGCGAGTAAATTTGCTACT 420
Db 361 CCAGTCTGATGTTGACTCTCGCGCGCGCAATCTCTGCGCGCGAGTAACTTAACTTAACT 420
QY 421 TCACCCCTGACATCTCTGTCGCTCTGCGCTCTGCGCTAAATTTAGTCTGATGACGCCCT 480
Db 421 TCTCCCTTACTTCTCGTGGCGCACCGGTACAAATTTGTTCTATACGCGCTCTCTCT 480
QY 481 AATCCGCTCTGCGCTGACGAGCTGACTTAATTAATTAATTAATTAATTAATTAATTA 540
Db 481 AGCCCACTTCTACCCCTCCAGAGCGGACCAATTAATTAATTAATTAATTAATTAATTA 540
QY 541 AATATGACAGTACCGGTTGCGCGCTACTATCCGTTTACCGCGCGCTAGTGCCTAAT 600
Db 541 AATATGCGCAGTACCGGTTGCTCGTGGCGCAATTCGCTACCGCGCGCTGCTCTCTCT 600
QY 601 GCAGTTGAGGCTATGCTATATCCATTTCTTCTGGCTCAAAACAAACAAACCCCTACA 660
Db 601 GCTGTTGGTGTACGCGCATCTCCATCTCTGTTCTGGCGCACAGACCAACCAACCCCGAG 660
QY 661 TCTGTTGACATGAATTCATTTACTTCCACTGATCTCAGGATTTCTGTTCAACCTGGCATA 720
Db 661 TCGGTTGACATGAATTCATTAACCTCAGCGATTTCTGTTATTTAGTCCAGCGCGGATA 720
QY 721 GCATCTGAATTTGCTCATCCCAAGCGAGCGCTTCTCACTACCGCAATCAAGTTTGGCGCT 780
Db 721 GCCTCGAGCTTGTATCCCAAGTGAGCGCTTACTACTACCGTAACCAAGTTTGGCGCTCT 780
QY 781 GTTGAGACATCTGTTGTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Db 781 GTTGAGACCTCCGGGTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
QY 841 CATGGCTCTCCAGTAACTCTCTATACCAATACCCCTTATACCGTGGCGCTTGGCTTACTG 900
Db 841 CATGGCTCACTGTAAATTTTATATATACACCTTATACCGTGGCGCTTGGCTTCTTG 900
QY 901 GACTTTGCGCTTAGAGCTTGTGTTTTCGCAATCTCACCACCTGTAAACCAATACACGTTG 960
Db 901 GACTTTGCGCTCGAATTTGAGTTCCGCACTCACCCCGGTAATACCAACACCGCGGTC 960
```


Db 241 CGACCACTCGGCTCGCTTGGCGTACACAGGCCACAGCCCGCGCTTGCCTCAAGTCGT 300
Qy 301 CGACCTGCACAGCCGGGCTCGCGCGCTGACGGCTGTGGCGCCCTGCCCATACACACTCA 360
Db 301 AGACCTACCAAGCTGGGGCCGCGCGCTAACCGGCTCGCTCCGGCCCATACACACCCG 360
Qy 361 CCGCTCCGGAGCTTGAATCTCGCGGTGCAATCTACCGCGCCAGTATATTTGCTACT 420
Db 361 CCAGTGCCTGATGTGCACTCCGCGCGGCGCACTTTGGCGCGGAGTATAAACCCTATCA 420
Qy 421 TCACCCCTGACATCTCTGTGGCCCTCTGGCACTAATTTAGTCTGTATGACGCCCCCTT 480
Db 421 TCTCCCTTACCTCTCTCGGTGCGCACCGGCACTAACTGGTTCTTTATGCGCCCTCTT 480
Qy 481 AATCGCCTCTGCCCTGACGAGCGGTACTAATCTACATTAATGCGCACAGAGCCCTCC 540
Db 481 AGTCGCTTTTACCCCTTCAGAGCGGCACTAATCCCATTAATGCGCACGAGCTTCT 540
Qy 541 AATTATGACAGTACCGGTTGCCCGCTACTATCCGTTACCGCCCTTAGTGCTTAAT 600
Db 541 AATTATGCCAGTACCGGTTGCCCGTGCACAAATCCGTTACCGCCCTGGTCCCAAT 600
Qy 601 GCAGTTGAGGCTATGCTATATCCATTTCTTGGCCCTCAAAACACCAACCCCTACA 660
Db 601 GCTGTGCGCGGTAGCGCATCTCCATCTCATTTCTGGCCACAGACCACACCCCGAG 660
Qy 661 TCTGTGACATGAATTCATTTACTTCCACTGATGTACGATTTCTTGTTCACCTGGCATA 720
Db 661 TCCGTTGATATGAATTCATTAACCTTCGACGAGATGTCGATTTAGTCCAGCCCGGATA 720
Qy 721 GCATCTGAATTTGTCATCCCAAGGAGCGCCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
Db 721 GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATGTAACCAAGCTGGCGCTCC 780
Qy 781 GTTCAGACATCTGTGTGCTGAGGAGGAGCGACCTCCGCTCTTGTCTATGTATGACATA 840
Db 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGCTGTGTATGCTTTGACATA 840
Qy 841 CATGGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTG 900
Db 841 CATGGCTACTCGTAATTTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 900
Qy 901 GACTTTGCTTAGAGCTTGAATTTGCAACCTTACCCCGGTTAACCAATACCGGGTCT 960
Db 901 GACTTTGCTTAGAGCTTGAATTTGCAACCTTACCCCGGTTAACCAATACCGGGTCT 960
Qy 961 TCCGTTTACTCCAGCACTGCTGCTCACTCC---GCCGAGGGCGGAGGAGCTGGGAG 1017
Db 961 TCCGTTTATTCAGCACTGCTGCGCACCGCCTTCTGTCGCGGTGGGAGGAGCTGGGAG 1020
Qy 1018 CTGACCAACTGACGACACCGAGGTTCTATGAAGATCTCCACTTTTACCGGCTTAATGGG 1077
Db 1021 CTCACCAACCGCTGTACCCGCTTTATGAAGAGCTCTATTTACTAGTACTAATGGT 1080
Qy 1078 GTAGTGAAGTGGCGGGGATAGCTCTAACAATTAACCTTGTGACACGCTCCTC 1137
Db 1081 GTGCGTGAGATCGCGCGGATAGCTCTCACTGTTCAACCTTGTGACACCTCTGCT 1140
Qy 1138 GCGGGCTCCGACAGATTAATTTGCTCGCTGGCGGCACTGTTTATTTCCGCGCG 1197
Db 1141 GCGGGCTCCGACAGATTAATTTGCTCGCTGGCGGCACTGTTTATTTCTACTCCCGTCC 1200
Qy 1198 GTTGTCTCAGCAATGGCGACCAACCGTGAAGCTCTATATCATCAGTGGAGATGCTCAG 1257
Db 1201 GTTGTCTCAGCAATGGCGACCACTGTTAAGTTGTATATCTGTAGAGATGCTCAG 1260
Qy 1258 CAGATTAAGGTGTGTATCCCCCAGATATCATCTTGGTGAATTCGCGTGGTGCATT 1317
Db 1261 CAGGATAAGGTATTTGCAATCCCGCATGACATTGACCTCGAGAACTCTCGTGGTTATT 1320
Qy 1318 CAGGATTTACACACAGATGACGAGATGCGGCCACCCCGTGGCTGGCCCATCTCGG 1377
Db 1321 CAGGATTTATTAACCAACATGAACAAGATCGGCGGAGCTTCTTCCAGCCCATCTCGC 1380

Qy 1378 CTTTTTCTGTCTCGAGCAATGATGTACTTTGGCTGTCCCTCACTGACGCCGAGTAT 1437
Db 1381 CTTTTTCTGTCTCGAGCTAATGATGTGCTTTGGCTCTCTCACTGCGCTGCGAGTAT 1440
Qy 1438 GACCACTCACCTTACGGGTGCTCAACTGCGCGGTTTATCTCGGACAGCGTGAATTG 1497
Db 1441 GACCACTCACCTTATGGCTCTTCTGACTGCGCCAGTTTATGTTCTGACTCTGTGACCTTG 1500
Qy 1498 GTGAATGTTGGAGTGGCGCGGAGCCGCTAGCCGATCGCTTGACTGTTGCTCAAAAGTCAC 1557
Db 1501 GTTAATGTTGGAGCGCGCGCGAGCCGCTGCGCGCTCGCTGATTGGACCAAGTCA 1560
Qy 1558 CTCGAGGGCGCGCCCTCCGACTGTTGAGCAATATTCCAAGACATTTCTTTGTGCTCC 1617
Db 1561 CTGACGGTGGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTCGCG 1620
Qy 1618 CTTCTGTGCAAGCTCTCCTTTTGGAGGCGCGCACAAACAAAGCAGGTTATCTTTATAAT 1677
Db 1621 CTCGCGGTAGCTCTCTTTCTGGAGGCGAGCACAACTAAAGCCGGGTACCTTTATAAT 1680
Qy 1678 TATTAATCTACTGCTAGTACAGATCTTGAATGCTGCGGCGCATCGGTCGCC 1737
Db 1681 TATAACACCACTGCTAGCGACCACTGCTTCTCGAGAATGCGCGGCGCACCGGTCGCT 1740
Qy 1738 ATTTCAACCTATACACAGGCTTGGGCGCGTCCGTCGCGCATTTCTGCGCGCGCGGTT 1797
Db 1741 ATTTCCACTTACACCACTAGCTGCGTGGTCTGCTGCTCCATTTCTGCGGTTGCCGCTT 1800
Qy 1798 TTGGCTCCAGCTCGCCCTGCTCTGCTGAGGATCTTTTGTGATTATCCGCGGCGGCG 1857
Db 1801 TTAGCCCCCACTCTGCTAGCATGCTTGGAGATACCTTGGACTACCTGCCCGCGCC 1860
Qy 1858 CACACATTTGATGATCTTCTGCCCTGAATGCCGCTTTAGGCTCCAGGCTTGTGCTTTC 1917
Db 1861 CATACTTTGATGATTTCTGCCAGAGTGCGCCCTTGCCCTTCAGGCTGCGCTTC 1920
Qy 1918 CAGTCAACTGCTGAGCTCCAGGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTG 1977
Db 1921 CAGTCACTGCTGAGCTTCAAGGCTTAAAGTTAAGTGGGTAAACCTCGGAGTTG 1980
Qy 1978 TAG 1980
Db 1981 TAG 1983

RESULT 9

US-10-165-868-3
; Sequence 3, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R.
Bradley, Daniel W.
Twu, Jr-Shin
Purdy, Michael A.
Tam, Albert W.
Krawczynski, Krzysztof Z.

TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/165,868


```
1561 CTTGACGGTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTCTTTGTCTCGCG 1620
1618 CTTGTGCAAGCTCTCTCTTTTGGAGCGCGCACAAACAAAGAGGTTATCTTATAAT 1677
1621 CTCGCGGTAAGCTCTCTTTCTGGAGGAGCAGCAACTAAAGCGGGTACCTTTATAAT 1680
1678 TATAATACTACTGTAGTGACAGACTTCTGATTGAAATGTCGCCGCCATCGGGTCGCC 1737
1681 TATAACACCACTGTAGGACCACTGCTTGTGAGATGCGCGGACCGGGTCTCCT 1740
1738 ATTTCACCTATACACAGAGGTTGGGCGCGTCCGTCGCCATTTCTGCGCGCGCGTT 1797
1741 ATTTCACCTATACACACTAGCTGGGTCTGGTCCCGTCCATTTCTGCGGTTGCCGTT 1800
1798 TTGGCTCCAGCTCGCCCTGGCTCTCTGAGGATACCTTTGATTATCCGGGGGGCG 1857
1801 TTAGCCCCCACTCTGGCTAGCATTTGCTTGGATACCTTTGGACTACCTGCGCGGCC 1860
1858 CACACATTTGATGACTTCTGCCCTGAATGCCGCGCTTTAGGCCCTCCAGGGTTGTCTTC 1917
1861 CATACTTTGATGATTTCTGCCAGAGTGCGCGCCCTTGGCTTCAGGGCTGCGCTTC 1920
1918 CAGTCACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTTG 1977
1921 CAGTCACTGTGCTGAGCTTCAGCGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
1978 TAG 1980
1981 TAG 1983
```

RESULT 10

```
US-10-165-868-1
; Sequence 1, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
;           Bradley, Daniel W.
;           Twu, Ji-Shin
;           Purdy, Michael A.
;           Tam, Albert W.
;           Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; FILING DATE: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
```

```
; FILING DATE: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2094 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BURMA SEQUENCE, FIGURE 7
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-165-868-1
```

Query Match

Best Local Similarity 69.1%; Score 1368.6; DB 15; Length 2094;
Matches 1606; Conservative 0; Mismatches 374; Indels 3; Gaps 1;

```
Qy 1 ATGCGCCTAGGCCTCTTTTGTCTGTTTCTCTCTTGTTCCTCTGTTCTGCTCTATGTTCCCGCGCA 60
Db 46 ATGCGCCTCGGCCTATTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 105
Qy 61 CCGACCCGTCAGCCGTCCTGCGCGCGCTGCTGGGGCGGCGAGCGCGGTACCGGGGGTGGT 120
Db 106 CCGCGCCGTCAGCCGTCCTGCGCGCGCTGCTGGGGCGGCGAGCGCGGTTCGGGGTGGT 165
Qy 121 TCTGCGGTCGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
Db 166 TCTGCGGTCGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 225
Qy 181 CCTTTGCCCGACAGCTTGCCTGCGCTGCGGGTCTGGAGCTCGCCCTTCGCCAACCAAGCC 240
Db 226 CCTTTGCCCGACAGCTTGCCTGCGCTGCGGGTCTGGAGCTCGCCCTTCGCCAACCAAGCC 285
Qy 241 CGGCACTTGGCTCCACTTGGCGAGATCAGGCCCGCGCCCTCGCTGCTGCTCCGCTGCG 300
Db 286 CGACCACTCGGCTCGCTTGGCGGTGACGAGCCCGCGCCCTCGCTGCTGCTCCGCTGCG 345
Qy 301 CGACCTGCCACAGCGGGGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 360
Db 346 AGACCTACACAGCTGGGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 405
Qy 361 CCGCTCCCGACGTTGATTCTCGCGGTGCAATTTACGCCGCCAGTATAATTTGCTACT 420
Db 406 CCAGTGCCTGATGTCGACTCCCGCGCGCATCTTGGCGCGGAGTATAACCTATCAACA 465
Qy 421 TCACCCCTGACATCCTCTGGGCTCTGGACATAATTTAGTCTCTGATGAGCGCCCTT 480
Db 466 TCTCCCTTACCTCTTCCGTGGCCACCGGCATAAACCCTGTTTATGCGCGCCCTCTT 525
Qy 481 AATCCGCTCTGCGCTGCGAGCGGTACTTAATCTACATTCACATTAATGCGCACAGAGGCTCC 540
Db 526 AGTCCGCTTTACCTCTTCCGAGCGGACCAATACCATATATATGCGCACAGAGGCTTCT 585
Qy 541 AATTATGCACAGTACCGGGTTCGCCGCTACTATCCGTTACCGGCCCTTAGTGCCTAAT 600
Db 586 AATTATGCCAGTACCGGGTTCGCCGCTGCGACAAATCCGTTACCGCGCTGTTCCCAAT 645
Qy 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTCTGGCTTCAAAACCAACCAACCCCTACA 660
Db 646 GCTGTGCGGGGTTACGCCATCTCCATCTATTCTGGCCACAGACCAACCAACCCCGAGC 705
Qy 661 TCTGTTGACATGAATTCCTACTTCTTCTGAGTGTGAGGATCTTCTTCAACTGCGATA 720
```

Db 706 TCCGTTGATGAAATCAATAAACCCTCGACGAGTGTTCGTATTTTAGTCCAGCCCGGCATA 765
Qy 721 GCATCTGAATTGGTCAATCCCAAGCAGCGCCTTCACTACCGCAATCAAGGTTGGGCTCG 780
Db 766 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCCTACACATATCGTAACCAAGGCTGGCGCTCC 825
Qy 781 GTTCGAGACATCTGGTGTGTGAGGAGGAAGCCACCTCCGCTCTGTGTCATGTTATGCATA 840
Db 826 GTCGAGACCTCTGGGTGGCTGAGGAGAGGTACCTCTGTCTGTGTTATGCTTTGCATA 885
Qy 841 CATGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 900
Db 886 CATGCTCACTCGTAAATTCCTATATAACACCTATACCGGTGCGCTTGGCTGTTG 945
Qy 901 GACTTTGCTTAGAGCTTGAGTTTCGGAATCTCACCACCTGTACACCAATACAGCTGTG 960
Db 946 GACTTTGCGCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC 1005
Qy 961 TCCGTTTACTCAGCACTGCTGTCACCTCC--GCCGAGGGGCCGAGCGGACTCGCGAG 1017
Db 1006 TCCGTTTATTCAGACATGCTCTCGCACCGCCTTCGTCGGGTGCGGACGGGACTGCGAG 1065
Qy 1018 CTGACCACACTGACGACCAACAGGTTTCAGAAAGATCTCCATTTACCGGCCCTTAATGGG 1077
Db 1066 CTCACCAACAGGCTGCTACCGCTTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1125
Qy 1078 GTAGTGAAGTCGCGCGGGTAGCTCTTAACATTACTTAACCTTGTGACACGCTCTCTC 1137
Db 1126 GTCGGTGAGATCGCGCGGGATAGCCCTCACCTGTTCACCTTGTGACACTCTGCTT 1185
Qy 1138 GCGCGGCTCCCGACAGAAATTAATTTGTCGGCTGCGGGCCAACTGTTTATTTCCGCGCG 1197
Db 1186 GCGCGCTTCGACAGAAATTTGTTTCGTCGGCTGTCGGCCAGCTGTCTACTCCCGTCC 1245
Qy 1198 GTTGTCTCAGCAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAACTCAG 1257
Db 1246 GTTGTCTCAGCAATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1305
Qy 1258 CAGGATAGGGTGTGCTATATCCCGACGATATCGATCTTGTGATTCGCGTGTGTGCTATT 1317
Db 1306 CAGGATAGGGTATTGCAATCCCGCATGACATTTGACCTCGAGAACTCTGTTGTGTTATT 1365
Qy 1318 CAGGATTTATGACAAACAGCATGAGCAGGATCGGCCACCCCGTCGCTCGCCATCTCGG 1377
Db 1366 CAGGATTTATGATAACCAACATGAACAGATCGCGAGCGCTTCTCCAGCCCATCGCGC 1425
Qy 1378 CTTTTTCTGTTCTCGAGCAATGATGTACTTTGCTGCTCTCTCTCAGCGCTCCGAGTAT 1437
Db 1426 CTTTTTCTGTTCTCGAGCTAATGATGTGCTTTGGCTCTCTCTCAGCGCTCCGAGTAT 1485
Qy 1438 GACCATGTCACCTTACGGGTCTCACTGCGCGGTTTATATCTCGACAGCGTACTTTG 1497
Db 1486 GACCATGTCACCTTATGGCTCTTCGACTGGCCGACGTTTATGTTCTGACTCTGTGACCTTG 1545
Qy 1498 GTGAATGTTGCGACTGCGCGCAGCGCGTAGCCGATCGCTTGACTGTGTCGCAAGTCAAC 1557
Db 1546 GTTAATGTTGCGACCGCGCGAGCGCGTTGCGCGTCTCGATTGGACCAAGTTCACA 1605
Qy 1558 CTCGAGCGGGCGGCTCTCCGACTGTGTAGCAATATTCAGACATTTCTTGTGCTCCCG 1617
Db 1606 CTTGAGCGGTGCGCGGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTGCGCG 1665
Qy 1618 CTTGTTGCGCAAGCTCTCTTTTGGGAGCGCGGCACAAACAAAGCAGTTATCTCTATAAT 1677
Db 1666 CTCGCGGTGAGCTCTCTTTCTGGGAGGAGGAGGACAACTAAGCCGGGTACCTTTATAT 1725
Qy 1678 TATAATTACTGCTAGTGAACAGATTTCTGATTTGAAAAATGCTGCGCGGCATCGGTGCGC 1737
Db 1726 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAATGCCCGCGGACCGCGGTGCT 1785
Qy 1738 ATTTCAACCTATACCAACAGCTTGGCGCGGTGCGGTGCGCATTTCTGCGCGCGGTT 1797
Db 1786 ATTTCCACTTACACCACTAGCCCTGGGTGCTGGTCCCGTCTCCATTTCTGCGGTGCGGTT 1845

Qy 1798 TTGGCTCCACGCTCGGCCCTGGCTCTGCTGAGGATACCTTTTGATTAATCCCGGGCGGCG 1857
Db 1846 TTAGCCCCCACTCTCGCTAGCATTTGCTTGAGGATACCTTTGGACTACCTCGCCGCGCC 1905
Qy 1858 CACACATTTGATGACTTCTGCCCTGAATGCGCGCTTTAGGCTCCAGGTTGTGCTTTC 1917
Db 1906 CATATTTGATGATTTCTGCCAGAGTGGCGCCCTTGGCTTCAGGGCTGCGCTTTC 1965
Qy 1918 CAGTCAACTGTCGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAAACTCGGAGTTG 1977
Db 1966 CAGTCTACTGCTGAGCTTCAGCGCTTAAGATGAAGTGGGTAAAACTCGGAGTTG 2025
Qy 1978 TAG 1980
Db 2026 TAG 2028

RESULT 11
US-09-851-410-6
; Sequence 6, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yarbough, Patrice O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; Fy, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,410
; FILING DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7195 base pairs
; TYPE: nucleic acid

Query Match	66.6%;	Score 1319.6;	DB 17;	Length 7256;
Best Local Similarity	79.5%;	Pred. No. 0;		
Matches 1575;	Conservative	0;	Mismatches 404;	Indels 3; Gaps 1;
QY	1	ATCGGCCCTAGGGCCTCTTTTGCCTGTTTCCTCTGTGTTTCGTGCTATGTTGCCCGCCCA	60	
Db	5190	ATCGGCCCTTAGGGCTGTTCTGTTGTTTCCTCTGTGCTTTTGCTCATGCTGCCCGCCA	5249	
QY	61	CCGACCGGTGAGCGCTCTGCGCGCGCTGTCGTGGGCGGCGACGCGCGGTACGGGGGTGGT	120	
Db	5250	CCGGCCGGTTCAGCCGTCTGCGCGCTGCGCGGTGGCGGCGAGCGCGGTGCGGGCGTGGT	5309	
QY	121	TTCGTGGGTGACCGGGTGTGATTCTCAGCCCTTCGCAAATCCCTATATTCATCAACCACAAC	180	
Db	5310	TTCTGGGGTGCACAGGGTGTGATTCTCAGCCCTTCGCGCTCCCTATATTCATCCAACCAAC	5369	
QY	181	CCCTTTTGCCCCACAGCTTGCGGCTGCGTCCGGGTCTGGACTCGCTCTTGCCCAACCAAGCC	240	
Db	5370	CCCTTCGCGCGCCGATGTCGTTTACAAACCGGGGCTGGAACTCGCCCTCGACAGCGCGCC	5429	
QY	241	CGGCCACTTGGCTCCACTTGGCGAGATCAGGCCCGACGCGCCCTCCGCTGCTCCCGTCGC	300	
Db	5430	CGCCCCCTTGGCTCCGCTTGGCTGGCGTGACCAATCCAGCGCCCTTCGCTGCGCCCGCCGCT	5489	
QY	301	CGACTTGCCACAGCGGGGGCTGCGGCGCTGACCGCTGTGGCGCTGCCCATGACACTCA	360	


```
Db 6217 CTTACTACACAGCAGCCACACGTTTCATGAGGACCTGCACTTCGCTGGCAGATGCG 6276
QY 1078 GTAGGTGAAGTCGGCGCGGAGTAGCTCTAAACATTACTTAACTTGCTGACAGCTTCCTC 1137
Db 6277 GTTGTGTAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6336
QY 1138 GCGGGGCTCCCGACAGAAATAATTTCTGCTGGCTGGCGGCAACTGTTTAAATCCCGCGG 1197
Db 6337 GCGGGTTTACCACAGAAATGATTTGCTGCGCTGGGGCAACTGTTTAAATCCCGCGG 6396
QY 1198 GTTGTCTAGCAATGGCGAGCAACCGTGAAGCTCTATACATCACTGAGGAATGCTCAG 1257
Db 6397 GTTGTCTAGCAATGGCGAGCAACAGTAAAGTTATATACATCTGTTGAGAATGCGCAG 6456
QY 1258 CAGGATAAGGGTGTGCTATCCCGACAGATATGATCTGTTGGTGAATCGCGTGTGTCATT 1317
Db 6457 CAGCAAGGGGATCACCATTCCATGATATAGACCTGGGTGATCCCGTGTGTTATC 6516
QY 1318 CAGGATATGACAAACAGCATGAGCAGATCGGCGCACCCCGCTGCGGCATCTCGG 1377
Db 6517 CAGGATATGATAACAGCAAGCAGCAGACCTACTCTCGTACCTGCCCTCTCGC 6576
QY 1378 CCTTTTCTGCTCCGAGCAATGATGATCTTTGGCTGTCCCTCACTGACCGAGTAT 1437
Db 6577 CCTTCTCAGTCTTCTGTCGCAATGATGTTTGTGCTTTCCCTCACTGCGCTGAGTAT 6636
QY 1438 GACCAGTCCACTTACGGGTGCTCACTGCGCGGTTTATATCTCGGAGGTGACTTTG 1497
Db 6637 GACCAGTACGATGCGGTGCTGCAACCAACCTATGATGCTCTGACACAGTTACGGTT 6696
QY 1498 GTGAATGTGCGACTGCGCGCAGCGCTAGCGCCGATCGCTGCTGCTGCTGCTGCTGCT 1557
Db 6697 GTTAATGTGGCTACTGTGCTCAGGCTGTGCGCGCTCCCTGATGCTTAAAGTTACT 6756
QY 1558 CTCGACGGCGCGGCTCCCGACTGTTGAGCAATTAATCCAGACATTTCTTTGTGCTCCCC 1617
Db 6757 CTGACCGCGCGCGGCTTACTTACCAATTCAGCAGTATCTAAGACATTTTATGTTCTCCG 6816
QY 1618 CTTCTGTGCAAGCTCTCTTTTGGGAGCGCGCACACAAAGCAGGTTATCCTTATAAT 1677
Db 6817 CTCGCGGGAAGCTGCTCTTTGGGAGGCTGGCAGCACTAAGCGCGGCTACCTTTACAA 6876
QY 1678 TATAATACTACTGTAGTACAGATCTGATTAATAATGCTGCGCGCATCGGTGCTGCC 1737
Db 6877 TATAATACTACTGTAGTACCAAAATTTGATGAGATGCGCGCGCACCGCTGCTGCT 6936
QY 1738 ATTTCACCTATACACAGGCTTGGGCGCGGTCCGGTTCGCAATTTCTGGGCGCGGTT 1797
Db 6937 ATTTCACCTATACACAGTATAGTTCGCGGCTCTTACCTCGATCTCTGCGGTGCGGTA 6996
QY 1798 TTGGCTCCAGCTCGCGGCTGCTGCTGAGGATCTTTGATTAATCCGGGCGGCGG 1857
Db 6997 CTGGCTCCACACTGCTGCGGCTGCTGCTGAGGATCTTTGAGGATCTATTGATTAACCGCGCGTGC 7056
QY 1858 CACACATTTGATGATCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1917
Db 7057 CATACTTTGATGATTTTTCGCGGAGTCCGCTACCTAGTTTTCAGAGGTTTTCATTC 7116
QY 1918 CACTCAACTGCTGCTGAGCTTCAGCGGCTTAAAGTTAAGTGGGTAAAACTCGGAGTTG 1977
Db 7117 CAGTCTACTATTGCTGAGCTCCAGCGTTTAAATAATGAAGTAGTAAAAACCGCGAGTCT 7176
QY 1978 TA 1979
Db 7177 TA 7178
```

RESULT 15

US-09-468-147-165

; Sequence 165, Application US/09468147A

; Publication No. US20030049601A1

; GENERAL INFORMATION:

```
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; EARLIER FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/173,141
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 165
; LENGTH: 7277
; TYPE: DNA
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: us2full
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (36)...(5159)
; OTHER INFORMATION: orf1
; NAME/KEY: CDS
; LOCATION: (5197)...(7176)
; OTHER INFORMATION: orf2
; OTHER INFORMATION: orf3 at positions 5159-5527
; OTHER INFORMATION: Xaa = Unknown or Other at position 322
; OTHER INFORMATION: Xaa = Unknown or Other at position 331
; OTHER INFORMATION: Xaa = Unknown or Other at position 445
; OTHER INFORMATION: Xaa = Unknown or Other at position 448
; OTHER INFORMATION: Xaa = Unknown or Other at position 634
; OTHER INFORMATION: Xaa = Unknown or Other at position 646
; OTHER INFORMATION: Xaa = Unknown or Other at position 811
; OTHER INFORMATION: Xaa = Unknown or Other at position 1553
; OTHER INFORMATION: Xaa = Unknown or Other at position 1578
; OTHER INFORMATION: Xaa = Unknown or Other at position 1691
; OTHER INFORMATION: Xaa = Unknown or Other at position 1792
; OTHER INFORMATION: Xaa = Unknown or Other at position 1938
; OTHER INFORMATION: Xaa = Unknown or Other at position 2155
; US-09-468-147-165
```

Query Match

Best Local Similarity 66.4%; Score 1315.2; DB 10; Length 7277;

Matches 1570; Conservative 3; Mismatches 406; Indels 3; Gaps 1;

```
QY 1 ATGCGCCTTAGGCTCTTTTGTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
Db 5197 ATGCGCCTTAGGCTCTTTCTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5256
QY 61 CCACCGGTGAGCGCTCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGT 120
Db 5257 CCGCGCGCGCAGCGCTCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGT 5316
QY 121 TCTGCGGTGACCGGTTGATTTCTGAGCCCTTCTGAGCTATCCCTATATTCATCAACCAAC 180
Db 5317 TCTGCGGTGACAGGTTGATTTCTGAGCCCTTCTGAGCCCTTCTGAGCTATATTCATCAACCAAC 5376
QY 181 CCCTTTGCCCGCAGACGTTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 240
Db 5377 CCCTTCCCGCGATGCTGCTTTTCAACACCGCGGCTGGAATCTGCGCTCGACAGCGCCC 5436
QY 241 CGGCCACTTGGTCTCACTTGGGAGATCAGGCGCCAGCGCCCTCCGCTGCGCTCCCGTCCG 300
Db 5437 CGCCCGCTTGGTCTCCGCTTGGGAGTACAGGCTCCAGCGCCCTCCGCTGCGCTCCCGTCCG 5496
QY 301 CGACCTGCCACAGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 360
```


Db	5497	CGATCTGCCCCAGCTGGGGCTGGCGCTGACTGCCGTGTCAACGGCTCCTGACACAGCC	5556
Qy	361	CCCGTCCCGGACGTTGATTTCTCGGGTGCAATTTCTACGCCGCCAGTAAATTTGTCTACT	420
Db	5557	CCTGTACTGATTTGACTACAGTGGTGCTATTTCTCGCGCGGACAGTACAAATTTGTCCACG	5616
Qy	421	TCACCCCTGACATCTCTGTGGCTCTGGCACTAAATTTAGTCCCTGATGACGCCCCCTTT	480
Db	5617	TCCCGCTCAGCTCATCTGTGCGTTTCGGGTACTAAATTTGGTCCCTATGTGCCCCGGTG	5676
Qy	481	AATCGCCTCTGCGCTCGAGACCGGTACTAATCTCACATTTATGGCCACACAGAGCCCTCC	540
Db	5677	AATCCCCCTTCTGCTCTCCAGATGCTACCACTCATATTTATGGCTACTGAGSCATCC	5736
Qy	541	AAATATGACAGTACCGGTTTCCCGCGTACTATTCGGTTACCGGCCCTTAGTGCTTAAT	600
Db	5737	AAATATGCCAGTATCGGTTTCTCGAGCTACAATCCGTTATCGGCCGCTGTGTCGAAT	5796
Qy	601	GCAGTTGGAGGTATGCTATATCCATTTCTTCTGGCCTCAAAACACACACCCCTACA	660
Db	5797	GCCGTTGGTGGTATGCGCAATTTCCATTTCTTTCTGGCCCCAACTACACTACCCCTACT	5856
Qy	661	TCGTGTGACATCAATTCATTTCACTTCACTGATGTCAGGATTTCTTTCAACTGGCATA	720
Db	5857	TCGTGATATGAATTTCTATTACTTCCACACATGATGTTAGGATTTGGTTTCAGCCCGGAT	5916
Qy	721	GCATCTGAATGGTCAATCCACAGCGAGCCCTTCACTACCGCAATCAAGGTTGGCGCTCG	780
Db	5917	GCTTCGAGCTAGTCAATCCCACTGAGCGCTTCAATACCGTAATCAAGGCTGGCGCTCT	5976
Qy	781	GTGAGACATCTGGTGTCTGAGGAGGAGCCACCTCGGCTCTGTCATGTTATGTCATA	840
Db	5977	GTGAGACACCGGTTGGCTGAGGAGGAGGCTACTTCCGCTGCTGTAATGCTTTGCATT	6036
Qy	841	CATGGCTCTCCAGTTAACTCTATACCAATFACCCCTTATACCGGTCCCTTGGCTTACTG	900
Db	6037	CATGGCTCTCTGTTAACTTCTTACACTAATACACCTTACACTGGTGGCGCTTCTT	6096
Qy	901	GACTTTTGCCTTAGAGTTGAGTTTCGGAATCTCACACCTGTAAACCAATACACGTG	960
Db	6097	GATTTTGCCTAGAGCTTGAATTTAGGAATTTGACACCCGGGAACCAACACCCCGTGT	6156
Qy	961	TCCCGTACTCCAGCACTGCTCTGCTAC---TCCGCCCGAGGGCCGACGGGACTGGGAG	1017
Db	6157	TCCCGGTATACCAAGCAGCCCGCCCGGCTCGCGCTGGTGTCTGATGGGACTGTGAG	6216
Qy	1018	CTGACCACAACTGACGACCACTGAGTTTCATGAAAGATCTCCACTTTACCGGCCCTTAATGG	1077
Db	6217	CTTTACTACCAAGCAGCCACACGTTTCATGAGGACCTGCACCTGCTGGCAGGAATGGC	6276
Qy	1078	GTAGGTGAAGTGGCGCGGGATGCTCTTAACATTTACTTAACCTTGTGTCACAGCTCCTC	1137
Db	6277	GTTGGTGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6336
Qy	1138	GGCGGGCTCCGACAGAAATTAATTTCTGGCTGGCGGCAACTGTTTATTTCCCGCCCG	1197
Db	6337	GGCGGTTTACGACAGAAATTAATTTCTGGCTGGCGGCAACTGTTTACTCCCGCCCG	6396
Qy	1198	GTTGTCTCAGCCAAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAG	1257
Db	6397	GTTGTCTCAGCCAAATGGCGAGCCAAACAGTAAAGTTATATACATCTGTTGAGAATGTCAG	6456
Qy	1258	CAGGATTAAGGTGTTGCTATCCCCACGATATCGATCTTGGTGAATTCGGTGGTGCATT	1317
Db	6457	CAAGACAAGGGCATCACCAATTTCCATGATATAGACTGGGTGACTCCCGTGGGTATC	6516
Qy	1318	CAGGATTAATGAACCAAGCATGACGATCGGCCCAACCGCTCGCTCGCTGGCCATCTCGG	1377
Db	6517	CAGGATTAATGAACCAAGCATGACGATCGGCCCAACCGCTCGCTCGCTGGCCCTCTCGC	6576
Qy	1378	CCTTTTCTGTTCTCCGAGCAATGATGTTTGGTGTCCCTCACTGACGCCGAGTAT	1437
Db	6577	CCCTTCTCAGTTCTTCTGTCGTCGAATGATGTTTGTGGCTTTCCCTCACTGCGCTGAGTAT	6636
Qy	1438	GACCAGTCCACTTACGGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCGTGAATTG	1497
Db	6637	GACCAGACTACGATGGGTCTCCACCAACCTATGATGCTCTCTGACACAGTTACGCTT	6696
Qy	1498	GTGAATGTTGGACCTGGCGCAGCCCTAGCCGATCGCTGACTGCTCCAAAGTCAACC	1557
Db	6697	GTTAATGTTGGCTACTGTTGCTCAGGCTTTGCCCGCTCCCTTGTGTTGCTAAAGTTACT	6756
Qy	1558	CTCGACGGGGGGCCCTCCCGACTGTTGAGCAATTTCCAAAGACATTTCTTTGTGCTCCCC	1617
Db	6757	CTGACGGCGCGCCCTTACTACCAATTCAGCAGTATTTCTAGACATTTTAIGTTCTCCCG	6816
Qy	1618	CTTCGTGCAAGCTCTCCTTTTGGGAGCCCGGCACAAACAAAGCAGGTTATCCTTATATAT	1677
Db	6817	CTCCGGGGAAGCTGCTCTTTTGGGAGGCTGGCAGCACTAAGGCCGGCTTACCCTTACAT	6876
Qy	1678	TATAATATCTACTGTAGTGACAGATTTCTGATTGAATGCTGCCGCCATCGGTCGCC	1737
Db	6877	TATAATATCTACCGCTAGTGACCAAAATTTTGAATTTGAGAAATGCGCGCCGCCACCGTGTGCT	6936
Qy	1738	ATTTCAACCTATACCAACAGGCTTGGGGCCGGTCCGGTCCGCAATTTCTGCGGCCCGGTT	1797
Db	6937	ATTTCCACTATACCACTAGCTTAGTGCCGGTCTTACCTCGATCTCTGCGGTGCGCGTA	6996
Qy	1798	TTGGCTCCACGCTCCGCGCTGCTCTGCTGAGGATATCTTTTGATTTATCCGGGGCGGCG	1857
Db	6997	CTGGCTCCACACTCTGCCCTTCTTGAGGATACTATTTGATTTACCCCGCCCGTGCC	7056
Qy	1858	CACATTTGATGACTTCTGCCCTGAATGCGCGCTTTAGGCTCCAGGGTTGTGCTTTC	1917
Db	7057	CATACTTTTGTGATTTTTCGCCGAGTGCCGTACCCCTAGGTTTGCAAGGTTGTGCAATTC	7116
Qy	1918	CAGTCAACTGCTGCTGAGCTCCAGCGCCCTTAAAGTTAAGTGGGTAAAACTCGGAGTTG	1977
Db	7117	CAGTCTACTATTGCTGAGCTCCAGCGTTTAAAAATGAAGTAGGTAAACCCCGGAGTCT	7176
Qy	1978	TA 1979	
Db	7177	TA 7178	

Search completed: August 21, 2004, 22:42:05
Job time : 893.329 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 13:02:51 ; Search time 57 Seconds
(without alignments)
3271.603 Million cell updates/sec

Title: US-09-851-410A-8

Perfect score: 3408

Sequence: 1 MRPRPILLLLMLPMLPAP.....QSTVAELQLKMKVGKTRRL 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3408	100.0	660	2	Aar38785 HEV ORF2
2	3408	100.0	660	2	Aar39306 Burma str
3	3408	100.0	660	2	Aar96089 Hepatitis
4	3408	100.0	660	2	Aaw35826 Hepatitis
5	3408	100.0	660	2	Aaw80197 Protein e
6	3408	100.0	660	2	Aaw93389 Human HEV
7	3408	100.0	660	3	Aab24120 Hepatitis
8	3408	100.0	660	4	Aab62523 HEV-Burma
9	3408	100.0	660	5	Aaol15699 Hepatitis
10	3408	100.0	660	7	AD024376 Hepatitis
11	3401	99.8	660	2	Aaw76369 Hepatitis
12	3401	99.8	660	2	Aaw71210 Protein e
13	3399	99.7	660	2	Aar70323 Hepatitis
14	3391	99.5	660	2	Aar91814 Hepatitis
15	3391	99.5	660	2	Aaw93386 Human HEV
16	3391	99.5	660	5	AB881670 Hepatitis
17	3391	99.5	660	6	ADA50060 Hepatitis
18	3391	99.5	660	7	ADB97807 HEV ORF2
19	3386	99.4	660	5	ABJ04803 Hepatitis
20	3383	99.3	660	4	AAE12380 Hepatitis
21	3382	99.2	660	2	AAE51265 HEV strai
22	3382	99.2	660	2	AAW81520 Hepatitis
23	3382	99.2	660	2	AAW93395 Human HEV
24	3382	99.2	660	2	AAW93388 Human HEV
25	3377	99.1	660	5	ABJ04811 Hepatitis

ALIGNMENTS

RESULT 1

AAR38785
ID AAR38785 standard; protein; 660 AA.

XX AAR38785;

XX AC

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 11-JAN-1994 (first entry)

XX HEV ORF2 protein.

DE Enterically transmitted non-A non-B hepatitis; vaccine.

XX Hepatitis E virus; Burma strain.

XX Key Location/Qualifiers

FT Peptide 225..660

FT Peptide /label= C2

FT Peptide 334..660

FT Peptide /label= SG3

FT Peptide 613..660

FT Peptide /label= 406.3-2

XX WO9314116-A1.

XX 22-JUL-1993.

XX 15-JAN-1993; 93WO-US000459.

XX 17-JAN-1992; 92US-00822335.

XX 01-MAY-1992; 92US-00876941.

XX (GENE-) GENELABS TECHNOLOGIES INC.

XX (USSH) US SEC DEPT HEALTH.

XX Reyes GR, Bradley DW, Tam AW, Carl M;

XX WPI; 1993-243144/30.

XX N-PSDB; AAQ47129.

XX New immunogenic hepatitis E virus (HEV) peptide(s) - are from the ORF1, ORF2 and ORF3 regions of HEV, useful as a vaccine against HEV infection.

XX Disclosure; Fig 8; 48pp; English.

XX Immunogenic hepatitis E virus (HEV) peptides are selected from the ORF1, ORF2 and ORF3 regions of HEV. The peptides can be used in vaccines to

26	3374	99.0	660	2	AAW93390	Human HEV
27	3372	98.9	660	2	AAW93394	Human HEV
28	3372	98.9	660	2	AAW93391	Human HEV
29	3370	98.9	660	5	ABJ04822	Hepatitis
30	3370	98.9	660	5	ABJ04819	Hepatitis
31	3369	98.9	660	5	ABJ04807	Hepatitis
32	3367	98.8	660	2	AAW93392	Human HEV
33	3363	98.7	660	5	ABJ04820	Hepatitis
34	3361	98.6	660	2	AAR14619	Protein e
35	3360	98.6	660	5	ABJ04806	Hepatitis
36	3360	98.6	660	5	ABJ04809	Hepatitis
37	3353	98.4	660	5	ABJ04821	Hepatitis
38	3352	98.4	660	5	ABJ04805	Hepatitis
39	3351	98.3	660	5	ABJ04813	Hepatitis
40	3351	98.3	660	5	ABJ04816	Hepatitis
41	3350	98.3	660	5	ABJ04810	Hepatitis
42	3349	98.3	660	5	ABJ04818	Hepatitis
43	3336	97.9	660	2	AAW93393	Human HEV
44	3290	96.5	655	5	ABJ04804	Hepatitis
45	3221	94.5	660	2	AAR96090	Hepatitis

CC prevent infection by HEV. The antibodies can neutralise and block HEV
CC infection and can be used to prevent or treat HEV infection. The peptides
CC and antibodies can also be used as diagnostic reagents. (Updated on 25-
CC MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS
CC field)

XX	Sequence 660 AA;	
XX	Query Match	100.0%; Score 3408; DB 2; Length 660;
XX	Best Local Similarity	100.0%; Pred. No. 5.3e-263;
XX	Matches 660; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MRPRPILLLLMLFPLPAPPQPGQRRRRGSGGGGFWGDRVDSQPPAIPYIHTN	60
Db	1 MRPRPILLLLMLFPLPAPPQPGQRRRRGSGGGGFWGDRVDSQPPAIPYIHTN	60
QY	61 PFAPDVTAAAGAGRVQPARPLGSARDQAPAVASRRRPTTAGAAPLTAVAPADHTP	120
Db	61 PFAPDVTAAAGAGRVQPARPLGSARDQAPAVASRRRPTTAGAAPLTAVAPADHTP	120
QY	121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTTHIMATEAS	180
Db	121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTTHIMATEAS	180
QY	181 NYAQYRVARATIRYRPLVNAVGYAISISFWPOTTTTPTSDVMNSITSDVRLVQPGI	240
Db	181 NYAQYRVARATIRYRPLVNAVGYAISISFWPOTTTTPTSDVMNSITSDVRLVQPGI	240
QY	241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSGLVMLCIHGSVNSYTNTPYTGALL	300
Db	241 ASELVIPSERLHYRNOGWRSVETSGVAEEATSGLVMLCIHGSVNSYTNTPYTGALL	300
QY	301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG	360
Db	301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG	360
QY	361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGPTVKLYTSVENAQ	420
Db	361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGPTVKLYTSVENAQ	420
QY	421 QDKGIALPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWLSLTAAY	480
Db	421 QDKGIALPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWLSLTAAY	480
QY	481 DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQYSKTFVLP	540
Db	481 DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQYSKTFVLP	540
QY	541 LRKGLSFWAGTTKAGYPYNNYNTTASDQLLVENAGHRVAISTYTSLGAGPVSISAVAV	600
Db	541 LRKGLSFWAGTTKAGYPYNNYNTTASDQLLVENAGHRVAISTYTSLGAGPVSISAVAV	600
QY	601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGTREL	660
Db	601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGTREL	660

RESULT 2
AAR39306
ID AAR39306 standard; protein; 660 AA.
XX AAR39306;
AC AAR39306;
XX
DT 25-MAR-2003 (revised)
DT 14-FEB-1994 (first entry)
XX
XX
DE Burma strain HEV ORF2 putative virus capsid protein.
XX
XX Hepatitis E virus; vaccine; neutralising antibodies; infection; block;
XX
XX open reading frame; antibodies.
OS
OS Hepatitis E virus.
XX

PN WO9314208-A2.
XX
PD 22-JUL-1993.
XX
PF 19-JAN-1993; 93WO-US000475.
XX
PR 17-JAN-1992; 92US-00822335.
PR 20-APR-1992; 92US-00870985.
XX
XX (GENE-) GENELABS TECHNOLOGIES INC.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Reyes GR, Bradley DW, Twu J, Purdy MA, Tam AW, Krawczynski KZ;
XX WFI; 1993-243223/30.
XX N-PSDB; AAQ46813.
XX
XX Antigen and antibody vaccines against hepatitis E virus infection -
XX contain peptide(s) derived from capsid protein C-terminal or antibodies
XX against protein.
XX
XX Disclosure; Fig 7; 43pp; English.
XX
XX The sequence is that of the putative virus capsid protein encoded by
XX Burma strain hepatitis E virus (HEV) open reading frame ORF2. This
XX protein or peptide fragments of it may be used in a vaccine composition
XX for immunising an individual against HEV. Antibodies raised against these
XX peptides can also be used in such vaccines. (Updated on 25-MAR-2003 to
XX correct PN field.)
XX
XX Sequence 660 AA;

Query Match	100.0.0%;	Score 3408;	DB 2;	Length 660;
Best Local Similarity	100.0.0%;	Pred. No. 5.3e-263;		
Matches 660;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MRPRPILLLLMLPMLPAPPQPGSGRRRRSGSGGGFWGDRVDSQPPAIPYIHTPN	60	
Db	1	MRPRPILLLLMLPMLPAPPQPGSGRRRRSGSGGGFWGDRVDSQPPAIPYIHTPN	60	
QY	61	PFAPDVTAAAGAGPRVQPARPLGSARWQAOQAPAVASRRRPTTAGAAPLTAVAPAHDP	120	
Db	61	PFAPDVTAAAGAGPRVQPARPLGSARWQAOQAPAVASRRRPTTAGAAPLTAVAPAHDP	120	
QY	121	PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTHIMATEAS	180	
Db	121	PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTHIMATEAS	180	
QY	181	NYAQYRVARATIRYRPLVNAVGGYAISSIFWPQTITPTSDVMNSITSDVRLVQPGI	240	
Db	181	NYAQYRVARATIRYRPLVNAVGGYAISSIFWPQTITPTSDVMNSITSDVRLVQPGI	240	
QY	241	ASELVIIPSERLHYRNOGWRSVETSGVAEEBATSGLVWMLCIHGSVNSYNTNPTYGALL	300	
Db	241	ASELVIIPSERLHYRNOGWRSVETSGVAEEBATSGLVWMLCIHGSVNSYNTNPTYGALL	300	
QY	301	DFALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG	360	
Db	301	DFALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG	360	
QY	361	VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSPRVVSANGEPPTVKLYTSVENAQ	420	
Db	361	VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSPRVVSANGEPPTVKLYTSVENAQ	420	
QY	421	QDKGIALPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWMLSLTAAEY	480	
Db	421	QDKGIALPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWMLSLTAAEY	480	
QY	481	DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQOYSKTFVLP	540	
Db	481	DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIQOYSKTFVLP	540	
QY	541	LRKGLSFWAGTTKAGYPYNTNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV	600	

```

Db 541 LRGLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAV 600
QY 601 LAPHSSALLLEDITLDYPARAHTFDPCPCPLGLOGCAFQSTVLAELORLKMVGKTREL 660
Db 601 LAPHSSALLLEDITLDYPARAHTFDPCPCPLGLOGCAFQSTVLAELORLKMVGKTREL 660

RESULT 3
AAR96089
ID AAR96089 standard; protein; 660 AA.
XX AC AAR96089;
XX DE Hepatitis E virus (Burma strain) capsid protein.
DT 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
XX KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
XX KW capsid.
XX OS Hepatitis E virus; Burma strain.
XX PN WO9612807-A2.
XX PD 02-MAY-1996.
XX PF 23-OCT-1995; 95WO-US013703.
XX PR 24-OCT-1994; 94US-00327952.
XX PR 13-OCT-1995; 95US-00542634.
XX PA (GENE-) GENELABS TECHNOLOGIES INC.
XX PI Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
XX DR WPI; 1996-230608/23.
XX DR N-PSDB; AAT27107.
XX PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
XX PT diagnostic reagents for determining HEV infection and in vaccines.
XX PS Disclosure; Page 78-80; 125pp; English.
XX CC The putative capsid protein (AAR96089) of hepatitis E virus (HEV) Burma
XX CC strain is encoded by ORF-2 (AAT27107) of the virus. PCR amplification of
XX CC ORF-2 allows prodn. of capsid protein or of C-terminal regions of the
XX CC capsid protein (see also AAR96091, AAR96093 and AAR96095) and expression
XX CC in Spodoptera frugiperda Sf9 insect cells using a baculovirus vector
XX CC provides recombinant C-terminal regions (see also AAR96101 and AAR96103)
XX CC useful as diagnostic reagents and in vaccines. The HEV Mexico strain
XX CC capsid protein (AAR96090) may similarly be used. (Updated on 16-OCT-2003
XX CC to standardise OS field)
XX SQ Sequence 660 AA;

Query Match 100.0%; Score 3408; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.e-263;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRFILLLLMFLPMLPAPPQSGRRGRSGSGGFGWDRVDSOPPAIPYIHTPN 60
Db 1 MRPRFILLLLMFLPMLPAPPQSGRRGRSGSGGFGWDRVDSOPPAIPYIHTPN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQQRPAVASRRRPTTAGAAPLTAVAPAHDT 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQQRPAVASRRRPTTAGAAPLTAVAPAHDT 120
QY 121 PVPDVSRCAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPDQGTNTHIMATEAS 180
Db 121 PVPDVSRCAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPDQGTNTHIMATEAS 180

```

```

QY 181 NYAQRYVARATIRYRPLVNAVGGYVAISISFPWPTTTPTSDVMNSITSTDVRLVQPGI 240
Db 181 NYAQRYVARATIRYRPLVNAVGGYVAISISFPWPTTTPTSDVMNSITSTDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNOQWRSVETSGVAEEATSLVMLCIHGSLVNSYNTNPTYGALGLL 300
Db 241 ASELVIPSERLHYRNOQWRSVETSGVAEEATSLVMLCIHGSLVNSYNTNPTYGALGLL 300
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLAADTLGGLPTELISSAGGOLFYSRPVVSANGBPTVKLYTSVNAQ 420
Db 361 VGEIGRGIALTLFNLAADTLGGLPTELISSAGGOLFYSRPVVSANGBPTVKLYTSVNAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHEQDRPTSPAPSRPESVLRANDVWLSTAAEY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHEQDRPTSPAPSRPESVLRANDVWLSTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSIDMTKVLTDGRPLSTIOQYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSIDMTKVLTDGRPLSTIOQYSKTFFVLP 540
QY 541 LRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAV 600
Db 541 LRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAV 600
QY 601 LAPHSSALLLEDITLDYPARAHTFDPCPCPLGLOGCAFQSTVLAELORLKMVGKTREL 660
Db 601 LAPHSSALLLEDITLDYPARAHTFDPCPCPLGLOGCAFQSTVLAELORLKMVGKTREL 660

RESULT 4
AAR35826
ID AAR35826 standard; protein; 660 AA.
XX AC AAR35826;
XX DT 17-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 26-FEB-1998 (first entry)
XX DE Hepatitis E virus Burma strain protein from ORF2.
XX KW Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody;
XX KW diagnosis; HEV.
XX OS Hepatitis E virus; - Burma strain.
XX PN US5686239-A.
XX PD 11-NOV-1997.
XX PF 09-MAY-1994; 94US-00240049.
XX PR 17-JUN-1988; 88US-00208997.
XX PR 11-APR-1989; 89US-00336672.
XX PR 16-JUN-1989; 89US-00367486.
XX PR 13-OCT-1989; 89US-00420921.
XX PR 05-APR-1990; 90US-00505888.
XX PR 05-APR-1991; 91US-00681078.
XX PR 17-JAN-1992; 92US-00822335.
XX PR 20-APR-1992; 92US-00870985.
XX PR 01-MAY-1992; 92US-00876941.
XX PA (GENE-) GENELABS TECHNOLOGIES INC.
XX PI Tam AW, Reyes GR, Yarbough PO;
XX DR WPI; 1997-558132/51.
XX DR N-PSDB; AAT96959.

```


QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNADTLGGLPTELISSAGGQLFYSPVPVVSANGEPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNADTLGGLPTELISSAGGQLFYSPVPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDLGESRVVIQDYDQHEQDRPTSPAPSPRPFVSLRANDVILWLSLTAAY 480
DB 421 QDKGIAIPHDIIDLGESRVVIQDYDQHEQDRPTSPAPSPRPFVSLRANDVILWLSLTAAY 480
QY 481 DQSTYSGSTGVPVVSVDVTLVNVATGAQAVARSIDWTKVTLDCRPLSTIQOYSKTFPVL 540
DB 481 DQSTYSGSTGVPVVSVDVTLVNVATGAQAVARSIDWTKVTLDCRPLSTIQOYSKTFPVL 540
QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
DB 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELTDLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660
DB 601 LAPHSALELTDLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660
RESULT 6
AAW93389
ID AAW93389 standard; protein; 660 AA.
AC AAW93389;
XX
XX
XX
XX 11-JUN-1999 (first entry)
XX
DE Human HEV ORF 2 protein from strain Burma.
XX
XX Swine hepatitis E virus; HEV; cross-reaction; antibody; human; therapy;
XX vaccine; immunise; infection; detection; diagnosis; prevention.
XX
OS Hepatitis E virus.
XX
XX WO9904029-A2.
PN
XX
XX 28-JAN-1999.
PF 17-JUL-1998; 98WO-US014665.
XX
PR 18-JUL-1997; 97US-0053069P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX Meng X, Emerson SU, Purcell RH;
XX
XX WPI; 1999-132270/11.
DR
XX
XX New isolated swine hepatitis E virus - used to develop products for the
PT diagnosis, prevention and treatment of hepatitis E virus infection in
PT mammals, particularly humans.
XX
XX Example 1; Fig 3A; 70pp; English.
XX
XX This invention describes a swine hepatitis E virus (HEV) and its natural
CC mutants which are capable of cross-reacting with antibodies reactive with
CC a human HEV strain or natural mutants. The HEV and the proteins can be
CC used in vaccines for immunising against HEV infection. The swine HEV can
CC be used in humans to prevent possible infection by human HEV. The swine
CC HEV can also be used as a therapeutic treatment for infection by other
CC strains of HEV. The swine HEV can also be used for the production of
CC antibodies which can be used in therapy, detection and diagnosis. The
CC products can also be used for determining the susceptibility of cells or
CC organs to infection with swine HEV. The swine HEV is particularly useful
CC for the development of agents for the prevention, treatment and detection
CC of human HEV because it is not a human virus and thus can be handled both

CC experimentally and clinically without fear of severe infection and/or
CC contamination
XX
SQ Sequence 660 AA;
Query Match 100.0%; Score 3408; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.3e-263;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMLPMLPAPPPQPSGRRRRSGSGGGFWGDRVDSQFPAIPIYHPTN 60
DB 1 MRPRPILLLLMLPMLPAPPPQPSGRRRRSGSGGGFWGDRVDSQFPAIPIYHPTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQAPAVASRRRPTTAGAAPTAVAPADHTP 120
DB 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQAPAVASRRRPTTAGAAPTAVAPADHTP 120
QY 121 PVPDVSRGAILLRQYNLSTPLTSSVATGTNLVLAAPLSPLLPDQDGNTHIMATEAS 180
DB 121 PVPDVSRGAILLRQYNLSTPLTSSVATGTNLVLAAPLSPLLPDQDGNTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAISSISFWPQTTTPTSVDMNSITSDVRLVQPGI 240
DB 181 NYAQYVARATIRYRPLVNAVGGYAISSISFWPQTTTPTSVDMNSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNQWRSVETSGVAEBEATSGLMVLCIHGSLVNSYNTPTGTALGLL 300
DB 241 ASELVIPSERLHYRNQWRSVETSGVAEBEATSGLMVLCIHGSLVNSYNTPTGTALGLL 300
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNADTLGGLPTELISSAGGQLFYSPVPVVSANGEPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNADTLGGLPTELISSAGGQLFYSPVPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDLGESRVVIQDYDQHEQDRPTSPAPSPRPFVSLRANDVILWLSLTAAY 480
DB 421 QDKGIAIPHDIIDLGESRVVIQDYDQHEQDRPTSPAPSPRPFVSLRANDVILWLSLTAAY 480
QY 481 DQSTYSGSTGVPVVSVDVTLVNVATGAQAVARSIDWTKVTLDCRPLSTIQOYSKTFPVL 540
DB 481 DQSTYSGSTGVPVVSVDVTLVNVATGAQAVARSIDWTKVTLDCRPLSTIQOYSKTFPVL 540
QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
DB 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELTDLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660
DB 601 LAPHSALELTDLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660
RESULT 7
AAB24120
ID AAB24120 standard; protein; 660 AA.
XX
XX AAB24120;
XX
XX 29-JAN-2001 (first entry)
XX
XX Hepatitis E virus Burma strain ORF2 protein sequence SEQ ID NO:8.
DE Hepatitis E virus; HEV; non-A/non-B viral hepatitis; viral protein;
CC immunological; diagnosis; hepatitis; infection; identification;
CC detection; immunoreactive; hepatotropic; antiinflammatory; virucide;
CC vaccine; antiviral; antigenic; antibody; antigen.
OS Hepatitis E virus.
XX
XX US6120988-A.
PN
XX

PD 19-SEP-2000.
 XX
 PF 07-JUN-1995; 95US-00478507.
 XX
 PR 17-JUN-1988; 88US-00208997.
 PR 11-APR-1989; 89US-00336672.
 PR 16-JUN-1989; 89US-00367486.
 PR 13-OCT-1989; 89US-00420921.
 PR 05-APR-1990; 90US-00505888.
 PR 05-APR-1991; 91US-00681078.
 PR 25-JUL-1994; 94US-00279823.
 XX
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PA (USGO) US GOVERNMENT.
 XX
 PI Yarbough PO, Krawczynski KZ, Fry KE, Bradley DW, Tam A, Reyes GR;
 DR WPI; 2000-593712/56.
 DR N-PSDB; AAA99259.
 XX
 PR Identifying recombinant antigen immunoreactive with antibody induced by
 PT hepatitis E virus (HEV), for detecting HEV infection, comprises
 PT immunoreacting a polypeptide from an HEV genome with an HEV-positive
 PT antiserum.
 XX
 PS Claim 6; Col 63-68; 46pp; English.
 XX
 CC The present invention describes a method for identifying a recombinant
 CC antigen immunoreactive with a hepatitis E virus (HEV)-induced antibody.
 CC The method comprises producing a polypeptide derived from an HEV genome,
 CC immunoreacting the polypeptide with an HEV-positive antiserum and
 CC selecting the polypeptide as a recombinant antigen if the polypeptide
 CC reacts with the HEV-positive antiserum. The method is useful for
 CC identifying recombinant antigen immunoreactive with antibody induced by
 CC HEV. The enterically transmitted non-A/non-B hepatitis (ET-NANB) (also
 CC known as HEV))-specific fragments are useful for identifying ET-NANB-
 CC derived cDNAs, which contain additional sequence information, as primers
 CC for detecting ET-NANB viral genomic material in a patient sample, for the
 CC synthesis of polypeptides for use in immunoassays, and for identifying
 CC similar antigenic regions encoded by related viral strains, e.g. Burmese
 CC strain. The antigens are especially useful in the preparation of vaccine
 CC against ET-NANB infection. These antigens may further be used to prepare
 CC antibodies to ET-NANB virus particles for use directly as antiviral
 CC agents, and to produce antiserum designed for pre- or post-exposure
 CC prophylaxis. The present sequence represents a specifically claimed HEV
 CC Burma strain protein sequence for use in the present invention
 XX
 SQ Sequence 660 AA;
 Query Match 100.0%; Score 3408; DB 3; Length 660;
 Best Local Similarity 100.0%; Pred. No. 5.3e-263; Indels 0; Gaps 0;
 Matches 660; Conservative 0; Mismatches 0;
 QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRRGSGGSGGFWGDRVDSQPPAIPYIHTN 60
 DB 1 MRPRPILLLLMLPMLPAPPQPSGRRRRGSGGSGGFWGDRVDSQPPAIPYIHTN 60
 QY 61 PFAPDVTAAAGAGRVQPARPLGSAWRDQORFAVASRRRPTTAGAPLTAVAPADTP 120
 DB 61 PFAPDVTAAAGAGRVQPARPLGSAWRDQORFAVASRRRPTTAGAPLTAVAPADTP 120
 QY 121 PVPDVSRGATLRQYNLSPLTSSVATGNLVLYAAPLSPLPLQDGTTHIMATEAS 180
 DB 121 PVPDVSRGATLRQYNLSPLTSSVATGNLVLYAAPLSPLPLQDGTTHIMATEAS 180
 QY 181 NYAQYVARATIRVPLVPNAVGGYASISFWPQTPTTPTSDVMNSITSTDRILVQPGI 240
 DB 181 NYAQYVARATIRVPLVPNAVGGYASISFWPQTPTTPTSDVMNSITSTDRILVQPGI 240
 QY 241 ASELVIPSERLHYRQGRSVETSGVAEEATSGIWMLCIHGSLVNSVTNTPYTGALGLL 300
 DB 241 ASELVIPSERLHYRQGRSVETSGVAEEATSGIWMLCIHGSLVNSVTNTPYTGALGLL 300

QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTAAATFRMKDLYFTSTNG 360
 DB 301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTAAATFRMKDLYFTSTNG 360
 QY 361 VGEIGRGIALTLFNADTLGGLPTELISSAGQQLFYSRPVVSANGEPTVKLYTSVENAQ 420
 DB 361 VGEIGRGIALTLFNADTLGGLPTELISSAGQQLFYSRPVVSANGEPTVKLYTSVENAQ 420
 QY 421 QDKGIALPHDIDIGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVWLMLSLTAAEY 480
 DB 421 QDKGIALPHDIDIGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVWLMLSLTAAEY 480
 QY 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVLDCRPLSTIQOYSKTFFVLP 540
 DB 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVLDCRPLSTIQOYSKTFFVLP 540
 QY 541 LRKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHRVAISTYTTSLGAGPVSISAVAV 600
 DB 541 LRKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHRVAISTYTTSLGAGPVSISAVAV 600
 QY 601 LAPHSALLALLEDTLDYPARAHTFDDFCPCRCPLGLOGCAFQSTVAELQRLKMKVGKTR 660
 DB 601 LAPHSALLALLEDTLDYPARAHTFDDFCPCRCPLGLOGCAFQSTVAELQRLKMKVGKTR 660
 RESULT 8
 AAB62523
 ID AAB62523 standard; protein; 660 AA.
 AC AAB62523;
 DT 23-JUL-2001 (first entry)
 DE HEV-Burma strain viral protein ORF2.
 KW HEV; enterically transmitted nonA/nonB viral hepatitis agent; viral;
 KW ET-NANB; infection; vaccine.
 OS Hepatitis E virus.
 XX US6223005-B1.
 XX 08-MAY-2001.
 PF 03-AUG-1998; 98US-00128275.
 PR 17-JUN-1988; 88US-00208997.
 PR 11-APR-1989; 89US-00336672.
 PR 16-JUN-1989; 89US-00367486.
 PR 13-OCT-1989; 89US-00420921.
 PR 05-APR-1990; 90US-00505888.
 PR 05-APR-1991; 91US-00681078.
 PR 25-JUL-1994; 94US-00279823.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 XX
 PI Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam AW;
 PI Fry KE;
 XX WPI; 2001-342705/36.
 DR N-PSDB; AAF83495.
 XX
 PT New DNA sequences of enterically transmitted non-A/non-B (ET-NANB)
 PT hepatitis viral agent, useful in diagnosing infection by an enterically
 PT transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.
 XX
 PS Disclosure; Col 63-68; 45pp; English.
 XX
 CC The invention relates to an isolated DNA comprising the genome of an
 CC enterically transmitted nonA/nonB (ET-NANB) viral hepatitis agent (also
 CC referred as HEV). The DNA sequences or their fragments are useful in
 CC preparing ET-NANB viral proteins and as probes for virus detection. These

CC	are particularly useful in diagnosing infection by an enterically	
CC	transmitted agent (e.g. ET-NANB virus), as well as in vaccine production.	
CC	The present sequence represents a ORF2 protein encoded by a ET-NANB viral	
CC	DNA sequence from HEV-Burma strain	
XX		
SQ	Sequence 660 AA;	
	Query Match	
	Best Local Similarity 100.0%; Score 3408; DB 4; Length 660;	
	Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MRPRILLLLMFLPMLPAPPQPSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60	
Db	1 MRPRILLLLMFLPMLPAPPQPSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60	
QY	61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHDT 120	
Db	61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHDT 120	
QY	121 PVPDVDSRGAILLRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEAS 180	
Db	121 PVPDVDSRGAILLRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEAS 180	
QY	181 NYAQYRVARATIRYRPLVPNAVGGVAISISFPQTTTPTSDVMNSITSTDVRLVQPGI 240	
Db	181 NYAQYRVARATIRYRPLVPNAVGGVAISISFPQTTTPTSDVMNSITSTDVRLVQPGI 240	
QY	241 ASELVIPISERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGALGLL 300	
Db	241 ASELVIPISERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGALGLL 300	
QY	301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360	
Db	301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360	
QY	361 VGEIGRGIALTLFNLADTLGLLPTELISSAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420	
Db	361 VGEIGRGIALTLFNLADTLGLLPTELISSAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420	
QY	421 QDKGIAIPHDDILGERSRVVIQDYNQHQDRPTSPAPSRPFSVLRANDVLMLSLTAAY 480	
Db	421 QDKGIAIPHDDILGERSRVVIQDYNQHQDRPTSPAPSRPFSVLRANDVLMLSLTAAY 480	
QY	481 DOSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTDGRPLSTIQYSKTFVLP 540	
Db	481 DOSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTDGRPLSTIQYSKTFVLP 540	
QY	541 LRKGLSFEAGTTKAGYPVNYNTTASDQLLVENAGHRAVLSYTTTSLGAGPVSISAV 600	
Db	541 LRKGLSFEAGTTKAGYPVNYNTTASDQLLVENAGHRAVLSYTTTSLGAGPVSISAV 600	
QY	601 LAPHASALALLEDTLPARAHTFDDFCECRPLGLGCAFOSTVAELQRKMKVGTREL 660	
Db	601 LAPHASALALLEDTLPARAHTFDDFCECRPLGLGCAFOSTVAELQRKMKVGTREL 660	
RESULT 9		
AAOI5699		
ID	AAOI5699 standard; protein; 660 AA.	
XX		
AC	AAOI5699;	
XX		
DT	08-NOV-2002 (first entry)	
XX		
DE	Hepatitis E virus (Burma strain) ORF2 protein.	
XX		
KW	HEV; enterically transmitted nonA/nonB hepatitis viral agent;	
XX	Burma strain; bile; ORF2.	
OS	Hepatitis E virus.	
XX		
PN	US6379891-B1.	
XX		

PD	30-APR-2002.	
XX		
PF	19-APR-2000; 2000US-00553427.	
XX		
PR	17-JUN-1988; 88US-00208997.	
PR	11-APR-1989; 89US-00336672.	
PR	16-JUN-1989; 89US-00367486.	
PR	13-OCT-1989; 89US-00420921.	
PR	05-APR-1990; 90US-00505888.	
PR	15-APR-1991; 91US-00681078.	
PR	25-JUL-1994; 94US-00279823.	
PR	07-JUN-1995; 95US-00478507.	
XX		
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PA	(GENE-) GENELABS TECHNOLOGIES INC.	
XX		
PI	Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam A, Fry KE;	
XX	N-PSDB; AAL50386.	
DR	WPI; 2002-517277/55.	
XX		
PT	Detecting the presence of enterically transmitted nonA/nonB hepatitis	
PT	viral (HEV) agents in bile samples from infected humans and monkeys using	
PT	polymerase chain reaction.	
XX		
PS	Disclosure; Col 19-24; 61pp; English.	
XX		
CC	The invention comprises a method for detecting the presence of	
CC	enterically transmitted nonA/nonB hepatitis viral (HEV) agents in a	
CC	sample and isolating HEV agents or nucleic acid fragments produced by the	
CC	agent. The method utilises PCR - using bile from a human or cynomolgus	
CC	monkey actively infected with HEV as a source of the agent. The method of	
CC	the invention is used for detecting the presence of a viral agent in a	
CC	sample of cultured cells infected with the agent and isolating	
CC	enterically transmitted nonA/nonB HEV agents or nucleic acid fragments	
CC	produced by the agent. The present amino acid sequence represents the	
CC	protein encoded by open reading frame 2 (ORF2) of a DNA sequence isolated	
CC	from a Burmese strain of the Hepatitis E virus	
XX		
SQ	Sequence 660 AA;	
	Query Match	
	Best Local Similarity 100.0%; Score 3408; DB 5; Length 660;	
	Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MRPRILLLLMFLPMLPAPPQPSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60	
Db	1 MRPRILLLLMFLPMLPAPPQPSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60	
QY	61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHDT 120	
Db	61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHDT 120	
QY	121 PVPDVDSRGAILLRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEAS 180	
Db	121 PVPDVDSRGAILLRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEAS 180	
QY	181 NYAQYRVARATIRYRPLVPNAVGGVAISISFPQTTTPTSDVMNSITSTDVRLVQPGI 240	
Db	181 NYAQYRVARATIRYRPLVPNAVGGVAISISFPQTTTPTSDVMNSITSTDVRLVQPGI 240	
QY	241 ASELVIPISERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGALGLL 300	
Db	241 ASELVIPISERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGALGLL 300	
QY	301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360	
Db	301 DFALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360	
QY	361 VGEIGRGIALTLFNLADTLGLLPTELISSAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420	
Db	361 VGEIGRGIALTLFNLADTLGLLPTELISSAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420	

QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVLDGRPLSTIQOYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVLDGRPLSTIQOYSKTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDLDYPARAHTFDDFCPCRPLGLQCAFQSTVAELQRLKMKVGTREL 660
Db 601 LAPHSALELLEDLDYPARAHTFDDFCPCRPLGLQCAFQSTVAELQRLKMKVGTREL 660

RESULT 10

ADD24376

ID ADD24376 standard; protein; 560 AA.

XX AC ADD24376;

DT 15-JAN-2004 (first entry)

DE Hepatitis E virus strain Burma ORF 2.

XX non-A/non-B viral hepatitis; ET-NANB; HEV; vaccine.

XX Hepatitis E virus; strain Burma.

PN US2003124510-A1.

XX 03-JUL-2003.

XX 07-MAY-2001; 2001US-00851410.

PR 17-JUN-1988; 88US-00208997.

PR 11-APR-1989; 89US-00336672.

PR 16-JUN-1989; 89US-00367486.

PR 13-OCT-1989; 89US-00420921.

PR 05-APR-1990; 90US-00505888.

PR 05-APR-1991; 91US-00681078.

PR 25-JUL-1994; 94US-00279823.

PR 03-AUG-1998; 98US-00128275.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam AW;

PI Fry KE;

DR WFI; 2003-810978/76.

DR N-PSDB; ADD24374.

XX New proteins derived from an enterically transmitted non-A/non-B (ET-NANB) viral hepatitis agent, useful for diagnosing ET-NANB infection and producing vaccines for immunizing an individual against ET-NANB hepatitis viral agent.

XX Disclosure; SEQ ID NO 8; 81pp; English.

XX The invention relates to a protein derived from an enterically transmitted non-A/non-B viral hepatitis agent. The viral proteins are useful for detecting infection by the enterically transmitted non-A/non-B (ET-NANB/HEV) hepatitis viral agent. The proteins are also useful for producing vaccines for immunising an individual against HEV hepatitis viral agent. The DNA fragments are useful for identifying and sequencing the entire viral agent and for assaying the viral agent in an infected sample. The present sequence is used in the exemplification of the invention.

XX Sequence 660 AA;

PA (DENK-) DENKA SEIKEN KK.

Query Match 100.0%; Score 3408; DB 7; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.3e-263;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMFLPMLPAPPQPSGRRRGRSGSGGFWGDRVDSQFPAIPYIHTPN 60
Db 1 MRPRPILLLLMFLPMLPAPPQPSGRRRGRSGSGGFWGDRVDSQFPAIPYIHTPN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSARWDOAQRPVAVASRRRPTTAGAAPTAVAPADHTP 120
Db 61 PFAPDVTAAAGAGPRVQPARPLGSARWDOAQRPVAVASRRRPTTAGAAPTAVAPADHTP 120
QY 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGGYAISISFWPQTPTTPTSDVMSITSDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGGYAISISFWPQTPTTPTSDVMSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEEATSLVWLCIHGSLVNSYTNTPYTGALL 300
Db 241 ASELVIPSERLHYRNOGWSRVETSGVAEEATSLVWLCIHGSLVNSYTNTPYTGALL 300
QY 301 DFALELEFRNLTPCNTNTRVSRYSSTARHRLRRGADGTAEITTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFRNLTPCNTNTRVSRYSSTARHRLRRGADGTAEITTAATRFMKDLYFTSTNG 360
QY 361 VGEIRGIALTLFNLADTLGLPLTBLISSAGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIRGIALTLFNLADTLGLPLTBLISSAGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVLDGRPLSTIQOYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDMTKVLDGRPLSTIQOYSKTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLLVENAAHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDLDYPARAHTFDDFCPCRPLGLQCAFQSTVAELQRLKMKVGTREL 660
Db 601 LAPHSALELLEDLDYPARAHTFDDFCPCRPLGLQCAFQSTVAELQRLKMKVGTREL 660

RESULT 11

AAW76369

ID AAW76369 standard; protein; 560 AA.

XX AC AAW76369;

DT 03-DEC-1998 (first entry)

DE Hepatitis E virus hollow particle protein #2.

XX Hollow particle protein; virus; antibody; detection; immunoassay;
XX infection.

XX Hepatitis virus.

XX JF10234383-A.

PD 08-SEP-1998.

XX 28-FEB-1997; 97JP-00062445.

XX 28-FEB-1997; 97JP-00062445.

PA (DENK-) DENKA SEIKEN KK.

Db	241	ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSVNSYNTPTYGALGLL	300
Qy	301	DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTAATRFMKOLYFTSTNG	360
Db	301	DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTAATRFMKOLYFTSTNG	360
Qy	361	VGEIGRGIALTLFNLADTLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ	420
Db	361	VGEIGRGIALTLFNLADTLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ	420
Qy	421	QDKGIALPHDIDIGESRVVIQDYDNOHEODRPTSPAPRPFSVLRANDVLMLSLTAAY	480
Db	421	QDKGIALPHDIDIGESRVVIQDYDNOHEODRPTSPAPRPFSVLRANDVLMLSLTAAY	480
Qy	481	DQSTYSGSTGPFVYVSDSVTLVNVATGAQAVARSLDWTKVTLDGRPLSTIQOYSKTFVLP	540
Db	481	DQSTYSGSTGPFVYVSDSVTLVNVATGAQAVARSLDWTKVTLDGRPLSTIQOYSKTFVLP	540
Qy	541	LRGKLSFWEAGTTKAGYPYNYNTTASDOLLVANAAGHRVAISTYTTSLGAGPVSISAVAV	600
Db	541	LRGKLSFWEAGTTKAGYPYNYNTTASDOLLVANAAGHRVAISTYTTSLGAGPVSISAVAV	600
Qy	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPGLQCAFSQSTVAELQRLKMKVGKTRREL	660
Db	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPGLQCAFSQSTVAELQRLKMKVGKTRREL	660
RESULT 13			
ID	AAR70323	standard; protein; 660 AA.	
XX	AAR70323;		
AC	XX		
DT	25-MAR-2003 (revised)		
DT	28-SEP-1995 (first entry)		
XX	Hepatitis E virus ORF2 protein.		
XX	HEV; ORF2; antigen; vaccine; immunogen.		
XX	Hepatitis E virus.		
XX	WO9508632-A1.		
XX	30-MAR-1995.		
XX	23-SEP-1994; 94WO-AU000572.		
XX	24-SEP-1993; 93AU-00001423.		
XX	15-DEC-1993; 93AU-00002964.		
XX	(MACF-) MACFARLANE BURNET CENT MEDICAL.		
XX	Anderson DA, Locarnini SA, Torresi J, Li F, Hui Z;		
XX	WPI; 1995-139601/18.		
XX	N-PSDB; AAQ86592.		
PT	Antigens of hepatitis E virus (HEV) - selectively immuno:reactive to		
PT	convalescent and/or acute phase circulating antibodies to HEV.		
XX	Disclosure; Page 39-41; 78pp; English.		
XX	RNA from an HEV strain isolated in the Xinjiang region of China was		
CC	subjected to RT-PCR to obtain fragments corres. to ORF2 (given in		
CC	AAQ86592) and ORF3 (AAQ86593) that encoded antigenic proteins (AAR70323-		
CC	24, respectively). DNA fragments were manipulated into pGEX vectors for		
CC	production of GST fusion proteins in E. coli. Applications include HEV		
CC	infection diagnosis, therapy and vaccine development. (Updated on 25-MAR-		
CC	2003 to correct PN field.)		
XX	Sequence 660 AA;		
SQ			

Query Match		99.7%; Score 3399; DB 2; Length 660;
Best Local Similarity		99.7%; Pred. No. 2.8e-262;
Matches 658; Conservative		1; Mismatches 1; Indels 0; Gaps 0;
Qy	1	MRPRPILLLLMLPMLPAPPFGSGRRRRGRRSGGGGFWGDRVDSQFFAIPYTHPTN 60
Db	1	MRPRPILLLLMLPMLPAPPFGSGRRRRGRRSGGGGFWGDRVDSQFFAIPYTHPTN 60
Qy	61	PFAPDVTAAGAGPRVQPARPLGSAWRDOAORPAVASRRRPTTAGAAPLTAAPAHDT 120
Db	61	PFAPDVTAAGAGPRVQPARPLGSAWRDOAORPAVASRRRPTTAGAAPLTAAPAHDT 120
Qy	121	PVPDVDSRGAILRRQYNLSTPLTSSVATGTLNLVYAAPLSPLPLQDGTNTHIMATEAS 180
Db	121	PVPDVDSRGAILRRQYNLSTPLTSSVATGTLNLVYAAPLSPLPLQDGTNTHIMATEAS 180
Qy	181	NYAQYVARATIRYRPLVNAVGGYAISSIFWQTTTTPTSDVMNSITSDVRLVQPGI 240
Db	181	NYAQYVARATIRYRPLVNAVGGYAISSIFWQTTTTPTSDVMNSITSDVRLVQPGI 240
Qy	241	ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSVNSYNTPTYGALGLL 300
Db	241	ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSVNSYNTPTYGALGLL 300
Qy	301	DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTAATRFMKDIYFTSTNG 360
Db	301	DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTAATRFMKDIYFTSTNG 360
Qy	361	VGEIGRGIALTLFNLADTLGGLPTELIISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db	361	VGEIGRGIALTLFNLADTLGGLPTELIISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
Qy	421	QDKGIALPHDIDIGESRVVIQDYDNOHEODRPTSPAPSPFVLRANDVLMLSLTAAY 480
Db	421	QDKGIALPHDIDIGESRVVIQDYDNOHEODRPTSPAPSPFVLRANDVLMLSLTAAY 480
Qy	481	DQSTYSGSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVTLDGRPLSTIQOYSKTFVLP 540
Db	481	DQSTYSGSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVTLDGRPLSTIQOYSKTFVLP 540
Qy	541	LRGKLSFWEAGTTKAGYPYNYNTTASDOLLVANAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db	541	LRGKLSFWEAGTTKAGYPYNYNTTASDOLLVANAAGHRVAISTYTTSLGAGPVSISAVAV 600
Qy	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPGLQCAFSQSTVAELQRLKMKVGKTRREL 660
Db	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPGLQCAFSQSTVAELQRLKMKVGKTRREL 660
RESULT 14		
AAR91814		
ID	AAR91814	standard; protein; 660 AA.
XX	XX	
AC	XX	
XX	XX	
DT	26-NOV-1996 (first entry)	
XX	XX	
DE	Hepatitis E virus strain SAR-55 ORF-2.	
XX	Hepatitis E virus.	
KW	Hepatitis E virus; HEV; SAR-55 strain; enteric transmission;	
KW	structural region; antigen; detection; antibody; vaccine; immunisation;	
KW	infection.	
XX	Hepatitis E virus.	
OS	XX	
XX	XX	
PN	WO9610580-A2.	
XX	XX	
PD	11-APR-1996.	
XX	XX	
XX	03-OCT-1995; 95WO-US013102.	
PF	XX	
XX	XX	
PR	03-OCT-1994; 94US-00316765.	
XX	XX	

(USSH) US DEPT HEALTH & HUMAN SERVICES.
Tsarev SA, Emerson SU, Purcell RH;
WPI: 1996-209320/21.
N-PSDB; AAT27394.
Isolated and purified hepatitis E virus strain SAR-55 DNA - encodes antigenic protein useful in diagnosis, prophylaxis and treatment of hepatitis E virus infection.
Claim 6; Page 13-15; 121pp; English.
The present sequence is the protein prod. of ORF-2 from the hepatitis E virus (HEV) strain SAR-55, which was implicated in an enterically transmitted non-A, non-B hepatitis in Pakistan. The protein encoded by the structural region of the virus (i.e. ORF-2), which is capable of forming HEV like particles, is useful for the detection of HEV antibodies (pref. IgG or IgM) in blood, plasma, sera, cerebrospinal fluid, tissue, urine or pleural fluid. The protein, and anti-HEV antibodies generated using the protein, can also be used in vaccines for immunising an animal against HEV infection. The protein is identified as a band of greater than 50 kD following SDS-PAGE of cell lysates of insect cells infected with a HEV ORF-2 conug. baculovirus, i.e. the claimed recombinant expression vectors pPIC9-1779, -1780 and -1781
Sequence 660 AA;
Query Match 99.5%; Score 3391; DB 2; Length 660;
Best Local Similarity 99.4%; Pred. No. 1.2e-261;
Matches 656; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 MRPRILLLLMLFPLPAPPQSGRRRRSGSGGFWGDRVDSOPFAIPYIHPTN 60
Db 1 MRPRILLLLMLFPLPAPPQSGRRRRSGSGGFWGDRVDSOPFAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAPAVASRRRTTAGAAPTAVAPADHTP 120
Db 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAPAVASRRRTTAGAAPTAVAPADHTP 120
QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRVPLVFNAGVYALSFWPQTTTTPTSDVMNSITSTDVRLVPGI 240
Query Match 99.5%; Score 3391; DB 2; Length 660;
Best Local Similarity 99.4%; Pred. No. 1.2e-261;
Matches 656; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 MRPRILLLLMLFPLPAPPQSGRRRRSGSGGFWGDRVDSOPFAIPYIHPTN 60
Db 1 MRPRILLLLMLFPLPAPPQSGRRRRSGSGGFWGDRVDSOPFAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAPAVASRRRTTAGAAPTAVAPADHTP 120
Db 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAPAVASRRRTTAGAAPTAVAPADHTP 120
QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRVPLVFNAGVYALSFWPQTTTTPTSDVMNSITSTDVRLVPGI 240

Db	181	NYAQVRVARATIRYRPLVFNAGGVAISISFWPQTTTTTTSVDMNSITSTDVRLVQPGL	240
Qy	241	ASELVIPSERLHYRNOGMSVETSGVAEEEAISGLVMLCIHGSLVNSYTNTPYTGAIGLL	300
Db	241	ASELVIPSERLHYRNOGMSVETSGVAEEEAISGLVMLCIHGSPVNSYTNTPYTGAIGLL	300
Qy	301	DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTTAAATRFMKDLYFTSTNG	360
Db	301	DFALELEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELETTTAAATRFMKDLYFTSTNG	360
Qy	361	VGEIGRGIALTLFNLADTLGGIPLTELISAGGQLFYSRPVVSANGEPVVKLYTSVENAQ	420
Db	361	VGEIGRGIALTLFNLADTLGGIPLTELISAGGQLFYSRPVVSANGEPVVKLYTSVENAQ	420
Qy	421	QDKGIAIPHDIDLGESRVVIQDYDNOHQDQDPTSPAPSRPFSVLRANDVLMLSLTAAEY	480
Db	421	QDKGIAIPHDIDLGESRVVIQDYDNOHQDQDPTSPAPSRPFSVLRANDVLMLSLTAAEY	480
Qy	481	DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLDCRPLSTIQOYSKTFVLP	540
Db	481	DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSIDWTKVTLDCRPLSTIQOYSKTFVLP	540
Qy	541	LRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV	600
Db	541	LRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV	600
Qy	601	LAPHSALELLEDLDYPARAHTFDDFCPCRPGLQGCAPQSTVAELQRLKMKVGTREL	660
Db	601	LAPHSVLALLEDTMDYPARAHTFDDFCPCRPGLQGCAPQSTVAELQRLKMKVGTREL	660

Search completed: August 16, 2004, 13:33:47
Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 13:28:02 ; Search time 20 Seconds
(without alignments)
3174.320 Million cell updates/sec

Title: US-09-851-410A-8

Perfect score: 3408

Sequence: 1 MRPRPILLLLMFLPMLPAP.....QSTVAFRLQRLMKVGTREL 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3408	100.0	660	1 VHHWH2	structural protein
2	3194.5	93.7	659	2 B44212	structural protein
3	143	4.2	1041	2 S55862	probable membrane
4	139.5	4.1	5188	2 B85547	probable RTX faml
5	139.5	4.1	5291	2 F90696	hypothetical prote
6	137	4.0	3570	2 T45025	mucin MUC5B, trach
7	135.5	4.0	1324	2 S52863	DNA-binding protei
8	135.5	4.0	3282	2 E82750	hemagglutinin-like
9	135.5	4.0	3442	2 E82589	hemagglutinin-like
10	135.5	4.0	3455	2 B82519	hemagglutinin-like
11	133	3.9	796	2 T21460	hypothetical prote
12	133	3.9	1802	2 S69703	HKR1 protein precu
13	132.5	3.9	583	2 S67571	hypothetical prote
14	130.5	3.8	1104	2 S59310	probable membrane
15	130.5	3.8	1161	2 S57180	probable membrane
16	130.5	3.8	1537	2 S33465	flocculation prote
17	130.5	3.8	3013	2 A80480	probable invasin y
18	129.5	3.8	635	2 F75477	hypothetical prote
19	129.5	3.8	1367	1 S48478	glucan 1,4-alpha-g
20	129.5	3.8	2130	2 A80821	probable exported
21	129	3.8	918	2 T02759	hypothetical prote
22	129	3.8	1015	2 J5062	phogrin precursor
23	128.5	3.8	1367	2 S1959	hypothetical prote
24	128	3.8	767	1 J00474	glucan 1,4-alpha-g
25	128	3.8	778	1 ALBYG	glucan 1,4-alpha-g
26	128	3.8	832	2 B87673	ABC transporter, H
27	127.5	3.7	518	1 P4WL47	L2 protein - human
28	127.5	3.7	800	2 T02852	probable membrane
29	127.5	3.7	860	2 T14650	tail fiber protein

ALIGNMENTS

RESULT 1

VHHWH2

structural protein 2 precursor - hepatitis E virus (strain Burma)

C:Species: hepatitis E virus

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 23-Jul-1999

C:Accession: C40778

R:Tam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.C.; Bradley, D.W.; Fry, K.E.; Reyes, G.

Virolgy 185, 120-131, 1991

A:Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length vi

A:Reference number: A40778; MUID:92024067; PMID:1926770

A:Accession: C40778

A:Molecule type: genomic RNA

A:Residues: 1-660 <TAM>

A:Cross-references: GB:M73218; NID:G330023; PID:AAA45736.1; PID:G330026

A:Note: the authors translated the codon CGC for residue 2 as Ala

C:Superfamily: hepatitis E virus structural protein 2

C:Keywords: structural protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-660/Product: structural protein 2 #status predicted <SP2>

Query Match 100.0%; Score 3408; DB 1; Length 660;

Best Local Similarity 100.0%; Pred. No. 8.2e-208; Indels 0; Gaps 0;

Matches 660; Conservative 0; Mismatches 0;

QY 1 MRPRPILLLLMFLPMLPAPPPGSGRRRSGSGGGGFWGDRVDSQPPFAIPYIHPTN 60

Db 1 MRPRPILLLLMFLPMLPAPPPGSGRRRSGSGGGGFWGDRVDSQPPFAIPYIHPTN 60

QY 61 PFAPDVTAAGAGPRVQPARPLGSAWDOQRPAVASRRRPTTAGAAPLTAVAPAHDT 120

Db 61 PFAPDVTAAGAGPRVQPARPLGSAWDOQRPAVASRRRPTTAGAAPLTAVAPAHDT 120

QY 121 PVPDVDSRGAILRQYNLSPLTSSVATGNLVLYAAPLPLPLQDGTNTHWATAS 180

Db 121 PVPDVDSRGAILRQYNLSPLTSSVATGNLVLYAAPLPLPLQDGTNTHWATAS 180

QY 181 NYAQVRVARATIRVPLVPNAVGVVAISISFWPQTTTPTSDVMNSITSTDVRLVQPGI 240

Db 181 NYAQVRVARATIRVPLVPNAVGVVAISISFWPQTTTPTSDVMNSITSTDVRLVQPGI 240

QY 241 ASELVIPSERLHYRNOQRVSVETSGVAEEATSGLWMLCIHGLSVNSVNTPTYGALGIL 300

Db 241 ASELVIPSERLHYRNOQRVSVETSGVAEEATSGLWMLCIHGLSVNSVNTPTYGALGIL 300

QY 301 DFALLEFRNLTPGNTNTRVSRYSSTAHRLRRGADGTAEITTTAAATFEMKDYFTSTNG 360

Db 301 DFALLEFRNLTPGNTNTRVSRYSSTAHRLRRGADGTAEITTTAAATFEMKDYFTSTNG 360

QY 361 VGEIGRGIALTLFNADTLGLGPLELISAGGOLFYSRPVVSANGEPYKLYTSVENAQ 420

Db 361 VGEIGRGIALTLFNADTLGLGPLELISAGGOLFYSRPVVSANGEPYKLYTSVENAQ 420

30 127.5 3.7 978 2 T14968
31 127 3.7 1275 2 T33369
32 127 3.7 7463 2 T36248
33 126.5 3.7 351 1 VCETC
34 126.5 3.7 948 2 T11678
35 126.5 3.7 1306 2 S25370
36 126 3.7 1037 2 T13350
37 125.5 3.7 725 2 A41258
38 125.5 3.7 1879 2 S74915
39 125.5 3.7 1952 2 T48814
40 124 3.6 963 2 T40290
41 124 3.6 1131 2 T41144
42 124 3.6 1483 2 C97012
43 123.5 3.6 797 1 VGBEX1
44 123 3.6 2232 2 T34434
45 122.5 3.6 1324 2 T14070

phage lambda-relat
hypothetical prote
CDA peptide synthe
coat protein - tur
hypothetical prote
MSB2 protein - yea
transcription fact
a-agglutinin core
extracellular nucl
hypothetical prote
hypothetical prote
hypothetical serin
probably celluloso
glycoprotein X pre
hypothetical prote
peptide synthetase

QY	421	QDKGIAIPHDIDLGESRVVIQYDNOHQDRPTTSPAPSRPFSVLRANDVWLSLTAAEY	480
Db	421	QDKGIAIPHDIDLGESRVVIQYDNOHQDRPTTSPAPSRPFSVLRANDVWLSLTAAEY	480
QY	481	DQSTYGSSTGPVYVDSVTLVNVATGAQAVARSLDWTKVTLDRPLSTLQQYSKTFEVL	540
Db	481	DQSTYGSSTGPVYVDSVTLVNVATGAQAVARSLDWTKVTLDRPLSTLQQYSKTFEVL	540
QY	541	LRGKLSFEWAGTTKAGYPYNYNTASDQLLVENAAGRHVAISTYTTSLGAGPVSISAVAV	600
Db	541	LRGKLSFEWAGTTKAGYPYNYNTASDQLLVENAAGRHVAISTYTTSLGAGPVSISAVAV	600
QY	601	LAPHSALALLEDTLDYPARAHTFDDCPCEPRLGQCAFQSTVAELQRLKMKVGKTREL	660
Db	601	LAPHSALALLEDTLDYPARAHTFDDCPCEPRLGQCAFQSTVAELQRLKMKVGKTREL	660

RESULT 2
B44212
structural protein 2 precursor - hepatitis E virus (strain Mexico)
C:Species: hepatitis E virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: B44212
R:Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.
A:Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HHV-191)
A:Reference number: A44212; MUID:93079857; PMID:1448913
A:Accession: B44212
A:Molecule type: Genomic RNA
A:Residues: 1-659 <GB>
A:Cross-references: HBW74506; NID:g330017; PID:AAA45732.1; PID:g330020
C:Superfamily: hepatitis E virus structural protein 2
C:Keywords: structural protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:31-659/Product: structural protein 2 #status predicted <SP2>

[illegible]

Db 480 DQSYTSGSGTPYVSDSVTLVNVATGAQAVARSILDMSKVTLDCGRPLPTVEQYSKTFVFLP 533

QY 541 LRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAAHRVAISYTTTSLGAGPVSISAVV 600

Db 540 LRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVENAAHRVAISYTTTSLGAGPVSISAAV 599

QY 601 LAPHSALALLEDTLDYPARAHTFDDFCEPRPLGLQCAFSQVLAELQRLKMKVGKTR 660

Db 600 LAPRSALALLEDTFDYDGRAHTFDDFCEPRALGLQCAFSQVLAELQRLKVKVGKTR 659

RESULT 3

SS5862

probable membrane protein YNL327w - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: hypothetical protein N0320

C;Species: *Saccharomyces cerevisiae*

C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000

C;Accession: S55862; S63308; S51288

R;Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.

Yeast 11, 567-572, 1995

A;Title: Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV ide

A;Reference number: S55859; MUID:95373280; PMID:7645347

A;Accession: S55862

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-1041 <MAF>

A;Cross-references: EMBL:Z46259; NID:G633655; PIDN:CAA86371.1; PID:G633659

R;Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.

submitted to the Protein Sequence Database, April 1996

A;Reference number: S63287

A;Accession: S63308

A;Molecule type: DNA

A;Residues: 1-1041 <MAW>

A;Cross-references: EMBL:Z71603; NID:G1302444; PID:e239572; PID:G1302445; MIPS

A;Experimental source: strain S288C

C;Genetics:

A;Gene: SGD:EGT2

A;Cross-references: SGD:S0005271; MIPS:YNL327w

A;Map position: 14L

C;Keywords: transmembrane protein

F;4-20/Domain: transmembrane #status predicted <TM1>

F;1025-1041/Domain: transmembrane #status predicted <TM2>

Query Match 4.2%; Score 143; DB 2; Length 1041;

Best Local Similarity 20.4%; Pred. No. 0.48;

Matches 144; Conservative 94; Mismatches 238; Indels 230; Gaps 310

QY 48 SQFATPYTHPNFADPVTAAGAGPRVQPARPLGSARWDQARPAVASRRRPT--- 104

Db 340 SQSSALPEV-----AVTASLSSG-----ILSSTVDGASTGADASMSAVSTVSS 382

QY 105 ---AGAAPLTAVAHADTPPVDPVDSRGAILRROYNLSLTPLTSVATGNLVLYAAPL 160

Db 383 SSEQASSSISLSAFSSN-----STFTTSPSSLSGATETYSIISSASI 425

QY 161 SPLPLQDGTNTHIMATEASNAVQRVAR---AIIRYPLVNAVGGVAISISFWPQTTT 217

Db 426 SVTQASYIDNSTTAVTQSTISTVASSAEKLSLSY-----TSNVTIISVSSATQHTT 478

QY 218 TPTSDVMNSTSDVRIILVQPGIASLVLIPSRHLHYRQWRSVETSGVAEEEAATSGLV 277

Db 479 TPTSVS-NSTTISSSSVL-----ESVISSEPLYLANTTVSG--- 511

QY 278 LCIHGLSVNSYNTNP-----TGALGLDLPFALEFRNLTPGNTNTRVRSYSTAR 322

Db 512 ---ASSASQSTNPYPVSVNSTTSATQIATAPFAINITGTISISSITNT--SSVSSTT- 566

QY 329 HRLRRG-----ADGTAEIATTAATRMKDL-----YFTSTNGVGEIGRG--- 367

Db 565 SSLSGGFFVSNNAVAGSVYILTTTETSAQITHGSLIPISITTTTSGTDKTSNKVA 621

QY 368 ---IALTLFNLADTLGGIPLTELSSA-----GGQL---FYSPRPVSVANGE 401

```
Db 625 SSTEIAQSVNNSSLSVSTINTNAATAAANARNATETHATSGSLQPSVHSSLLSSTID 684
QY 408 PTVKLYTVENAOQDQKIAIPHDIDIGESRVVLQDYDNQHEQDRPFPSPFVSLRA 467
Db 685 --TKVTIATSTSRDGSLSLAFTTGLNQSVVTGTD-----KSDTYSVIS 727
QY 468 NDVLMISLTAAYEQD---STY-----GSSTGPVYVSDSVTLVNVATG-----AOAVAR 512
Db 728 TE-----SAQVTEYDSLPLSTLKPTVVTGTSRNSTFSMSSTKLTEATATDKGDAYSVLIS 783
QY 513 SLQWTKVTLQDGR--PLSTIQ-----QYSKTFEVLPLRGK 544
Db 784 STQSAQVTEYSGMLSPISLTLETPTVIMSTDESGYFTLTTCESQATEYGLSIPISLTDGS 843
QY 545 LSPWEAG-----TTKAGYPYNYNTASDQLLVENAAHRAHVAIST-----YTSILG 589
Db 844 VIYFTGESVVGYSYTVGAAQIAQHTSL---VPVSTIKGSKTSLSTESVWAGYSTTVG 900
QY 590 AG-----PVSI---SAVAVLAPHLSALALLEDTLDYPARA-HT 622
Db 901 AAQYAQHTSIVPVSTIKGSKTSLSTESVVGYSYTVGSAQYAHT 946

RESULT 4
probable RTX family exoprotein [imported] - Escherichia coli (strain O157:H7, substrain
B85547
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: B85547
R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85547
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5188 <STO>
A:Cross-references: GB:AE005174; NID:gl2513368; PIDN:AA054838.1; GSPDB:GN00145; UWGP:Z06
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z0615

Query Match 4.1%; Score 139.5; DB 2; Length 5188;
Best Local Similarity 22.3%; Pred. No. 8;
Matches 131; Conservative 70; Mismatches 235; Indels 151; Gaps 28;

QY 60 NPFAPDVTAAAGAGPRVRQPARPLGSW-----RDOAQRPAVASRRRPTTAGAA 108
Db 2478 NSVAYSATVQADGWSVVGIPAAV-SAMPAGPLTVEVDGQSSANNPVSVS--HPPT---V 2531
QY 109 PLTAVAPADHTPPVDVDSRGAILRRQVNLSTPSLTSSVATGTLNLVLAAPLSPLLPQD 168
Db 2532 DLTAVAISINTVASDDVINAA---BKGTNLTLGSGTSGIESQTVT-----VTF 2577
QY 169 GNTHIMATEASNYAQYVARATIRYP-----LYPNAVGGYALISIFWPQTITPT 220
Db 2578 GSKTYSVAANGSVSNVPAADLATLPEGAANVOASVSSAGNSASATHATSVSDASPT 2637
QY 221 SVDNMSITSDVRILVQPGIASSELVIPSLRHYRQGRSVETSGVABEEATSGLVMLCI 280
Db 2638 -LTINTIASDDILNAAEAG--SPLTI-----SGTSTAE-TGQITVTVL 2676
QY 281 HGLSVNSTNTPY-----TGALGLDFALELEFRNLTPGNTNTRVSRYSSTARH 330
Db 2677 NGA---TYTGTVQADGWSVSVPTSALGALNAS-----NYTVSATVNDKAGNPGSASHN 2727
QY 331 LRRGADGTA---ELTTTAAATRFMKD-----LYFTSTNGVGEIGRGIALTLNLAADTLG 381
Db 2728 L--AVDTTAPVLTINTVAGDDIINDAEHAQALVISTSGGEXGDVSWVL----- 2776
QY 382 GLPTELISAGQQLFYSPVPSVANGEPTVKLYTSVENAQDQKIAIP-HDID-LGESRV 439
Db 2777 -----NG-----KYYTTILDAGNWSVGVPAADVTALGSGAQT 2809
QY 440 IQDYDNQHEQDRPFPSPFVSLRANDVLMISLTAAB-----YDQSTYGSSTG 490
Db 2810 I---TASVSDRAGNSDDASRTVTVLSAPVISINTIAGDDVINATEKGSDLALSGTSDQ 2865
QY 491 PVYVSDSVTLNVNATGAQAVARSLDWTKVTLQDGRPLSTIQO--YSKTFEVLPLRGKLSF- 547
```

```
Db 2777 -----NG-----KYYTTILDAGNWSVGVPAADVTALGSGAQT 2809
QY 440 IQDYDNQHEQDRPFPSPFVSLRANDVLMISLTAAB-----YDQSTYGSSTG 490
Db 2810 I---TASVSDRAGNSDDASRTVTVLSAPVISINTIAGDDVINATEKGSDLALSGTSDQ 2865
QY 491 PVYVSDSVTLNVNATGAQAVARSLDWTKVTLQDGRPLSTIQO--YSKTFEVLPLRGKLSF- 547
Db 2866 PAGTALTIVTL-NGQNVYSATTDASGNWS-VTVPASAVSALGEATYSVTASVTNAQGNSSTA 2923
QY 548 --WEAGTTKAGYPYNYNTASDQLLVENAAHRAHVAISTVYTSILGAG 591
Db 2924 SHNVQNTALPG--ITINPVATDDIINASEAGSAQTISGQVTCGAAG 2968

RESULT 5
F90696
hypothetical protein ECs0542 [imported] - Escherichia coli (strain O157:H7, substrain RI
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: F90696
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F90696
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5291 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA033965.1; PID:gl3360000; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs0542

Query Match 4.1%; Score 139.5; DB 2; Length 5291;
Best Local Similarity 22.3%; Pred. No. 8.2;
Matches 131; Conservative 70; Mismatches 235; Indels 151; Gaps 28;

QY 60 NPFAPDVTAAAGAGPRVRQPARPLGSW-----RDOAQRPAVASRRRPTTAGAA 108
Db 2478 NSVAYSATVQADGWSVVGIPAAV-SAMPAGPLTVEVDGQSSANNPVSVS--HPPT---V 2531
QY 109 PLTAVAPADHTPPVDVDSRGAILRRQVNLSTPSLTSSVATGTLNLVLAAPLSPLLPQD 168
Db 2532 DLTAVAISINTVASDDVINAA---BKGTNLTLGSGTSGIESQTVT-----VTF 2577
QY 169 GNTHIMATEASNYAQYVARATIRYP-----LYPNAVGGYALISIFWPQTITPT 220
Db 2578 GSKTYSVAANGSVSNVPAADLATLPEGAANVOASVSSAGNSASATHATSVSDASPT 2637
QY 221 SVDNMSITSDVRILVQPGIASSELVIPSLRHYRQGRSVETSGVABEEATSGLVMLCI 280
Db 2638 -LTINTIASDDILNAAEAG--SPLTI-----SGTSTAE-TGQITVTVL 2676
QY 281 HGLSVNSTNTPY-----TGALGLDFALELEFRNLTPGNTNTRVSRYSSTARH 330
Db 2677 NGA---TYTGTVQADGWSVSVPTSALGALNAS-----NYTVSATVNDKAGNPGSASHN 2727
QY 331 LRRGADGTA---ELTTTAAATRFMKD-----LYFTSTNGVGEIGRGIALTLNLAADTLG 381
Db 2728 L--AVDTTAPVLTINTVAGDDIINDAEHAQALVISTSGGAGDGVSWVL----- 2776
QY 382 GLPTELISAGQQLFYSPVPSVANGEPTVKLYTSVENAQDQKIAIP-HDID-LGESRV 439
Db 2777 -----NG-----KYYTTILDAGNWSVGVPAADVTALGSGAQT 2809
QY 440 IQDYDNQHEQDRPFPSPFVSLRANDVLMISLTAAB-----YDQSTYGSSTG 490
Db 2810 I---TASVSDRAGNSDDASRTVTVLSAPVISINTIAGDDVINATEKGSDLALSGTSDQ 2865
QY 491 PVYVSDSVTLNVNATGAQAVARSLDWTKVTLQDGRPLSTIQO--YSKTFEVLPLRGKLSF- 547
```

Db 2866 PAGTALTIVTL-NGQNSVATTDASGNWS-VTVFASVASLGEATYSVTASVTHAQNSSSTA 2923

Qy 548 ---WEAGTTKAGYPYNYNTTASDQLAVENAAGHRVAISTYTTSLGAG 591

Db 2924 SHNVQVNTALPG--ITINPVATDDIINASEAGSAQTISQVTTGAAG 2968

RESULT 6

T45025

mucin MUC5B, tracheobronchial [imported] - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C:Accession: T45025

R:Desseyn, J.L.; Guyonnet-Duperat, V.; Porchet, N.; Aubert, J.P.; Laine, A. J. Biol. Chem. 272, 3168-3178, 1997

A:Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alternat

A:Reference number: Z22899; MUID:97166151; PMID:9013550

A:Accession: T45025

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-3570 <DES>

A:Cross-references: EMBL:Z72496; NID:g1834502; PIDN:CAA96577.1; PID:g1834503

A:Experimental source: placenta

C:Genetics:

A:Gene: MUC5B

Query Match 4.0%; Score 137; DB 2; Length 3570;

Best Local Similarity 20.1%; Pred. No. 6.7;

Matches 143; Conservative 72; Mismatches 257; Indels 240; Gaps 33;

Qy 52 AIPYIH--PNNPAPDVTAAGAGPRVQPARPLGSNWRDQORPAVASRRRTTAGAAP 109

Db 1867 ALPALRSTATTATSVTAI-----PSSSLGTAW-----TSLQTTTPTAT 1907

Qy 110 LTAVAPAHDTPPVPDVSRGAILLRQYNLSTPLTSSVATGTMNLVYAAPLSPLLPDQG 169

Db 1908 MSTATPS-STPE-----TVHTSVLTATTTRTGSVATPS-----TPG 1946

Qy 170 T--NTHIMATEASNYAQYRVARATIRYRPLVFNAGVYASISFWPQTTTTPT 220

Db 1947 TAHTTKVPTTTTGFT-----ATPSSSPGTALTTPPWISITTTTPTTRGSTVT 1993

Qy 221 --SVDMSITSDVRIIVQGIASELVIPIERLHYRQGRSVETSGVAEEATSGLVML 278

Db 1994 PSSIPGTHTATVLTITTTTATVATGSMATPSS-----STQTSCTPSSLTTATIT 2043

Qy 279 CIHGSLVN-----SYTNTPYTGALG-----LLDFALELEF 308

Db 2044 AT-GSTNPSSTGCTPIPPVLITATTATPAATSSVTTP-SSALGTHTPPVPNTATTGG 2101

Qy 309 RNTTPGNTNTRVSRYSSTARHLRRLRGADGTAELT-----TTAAT----- 347

Db 2102 RSLPPSPHTVPTAWTSAT-----SGILGTHITEPSTGTSHTPAATGTTQBPSTPALSS 2156

Qy 348 ----RPMKDLFTSTNGVGEICRGIALTLFNL-----ADTLGLGPTE-----LISSAG 392

Db 2157 PHPSSRTSPSPSGTTTPGHT-RGTSRTTATATPSTRKTSILLPSSPTSAPITVVITG 2215

Qy 393 GQ-----LFYSRPVVSANGEPVTKLYTSVE-----NAQQDKGI----- 425

Db 2216 CEPQCAWSEMLDYSPMPGPGSG-GDFETYSNIRAAGGAVCEQPLGLECRAQAQPGVPLRE 2274

Qy 426 ---AIPHDIDL-----ESRVVIQDYNQHQEDRP-----TPSPAP 458

Db 2275 LGQVVECSLDFGLVCRNREQVGKFCMCFNVEIRVFCFCNYG--HCPSTPATSSATPSSSTP 2332

Qy 459 SRPESVLRANDVLMLSLTAAYD-----QSTYGSSTGPVYVSDVTLVNVATGAQAV 510

Db 2333 GTTWILTK-----LTTATTTESTGSTATPSSTQGGPAGTPHVTATTPTV-TSSKAT 2385

Qy 511 ARSLDWTKVLIDGRPLSTIQOYSKTFVFLPLRGLKLSFWEAGTTKAGYPYNYNTTASDQL 570

Db 2386 PFSSPGTATAPALRSTATTATPTATSPATFSPSSSLGTTW-TLSQTTTPMATMSTATPSSST 2444

Qy 571 VENAAGHRVAISTYTTSLGAGPVVISAVAVLAPHASLALLEDTLDVYPARAHT 622

Db 2445 PETVHTSTVLTTATTATTGATG-----SVATPSSST-----PGTAHT 2479

RESULT 7

S52863

DNA-binding protein R kappa B - human

C:Species: Homo sapiens (man)

C>Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 01-Dec-2000

C:Accession: S52863; A45580

R:Nieters, A.; Boumeester, T.; Scheidereit, C. submitted to the EMBL Data Library, August 1994

A:Reference number: S52863

A:Accession: S52863

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1324 <NIE>

A:Cross-references: EMBL:X80878; NID:g695578; PID:g695579

R:Adams, B.S.; Leung, K.Y.; Hanley, E.W.; Nabel, G.J. New Biol. 3, 1063-1073, 1991

A:Title: Cloning of R kappa B, a novel DNA-binding protein that recognizes the interleuk

A:Reference number: A45580; MUID:92135142; PMID:1777480

A:Accession: A45580

A:Molecule type: mRNA

A:Residues: 313-1292, 'ISNREQLIAPP', 1304, 'QSPLLL', <ADA>

A:Cross-references: GB:U08191; GB:S79520; NID:g476273; PIDN:AAA1781.1; PID:g476274

A:Note: sequence extracted from NCBI backbone (NCBIN:79520)

C:Keywords: DNA binding; transcription factor

Query Match 4.0%; Score 135.5; DB 2; Length 1324;

Best Local Similarity 19.0%; Pred. No. 2;

Matches 131; Conservative 95; Mismatches 232; Indels 231; Gaps 29;

Qy 68 AAAGAGPRVRQARP-----LGSARWDQORPAVASRRRT-----TA 105

Db 693 AAARKALKQKPKPKSVKSSKSSIKVLSSGPSEQSOMSLSDSSMPTTPVTPPTT 752

Qy 106 GAAPLTAVAPAHDTPPVPDVSRGAILLRQYNLSTPLTSSVATGTMNLVYAAPLSPL- 164

Db 753 PALPALPISP-----PPVSAVNKSG-----PSTVSEPAKSSGVLLVSSPTMHLG 798

Qy 165 -----PLQDGTNTHIMATEASNYAQYR-VARATIRYRPLVFNAGVYASIS- 209

Db 799 TMLSPASSQTAPSSQAARVSHSGSAGLSQVRVVAQPSL---PAVQQSGGPAQTLPQM 855

Qy 210 -----SFWPQTT-----TTPTSVDMSITSDVRIIVQGIASELV 245

Db 856 PAGQIRVPATATQTKVVPQTMATVPVKAQTAAATVQRPQGQTLTSLPATASPV 915

Qy 246 IPIERLHYRQGRSVETSGVAEEATSGLVMLCIHGLSVNSYNTPTTGALG- 299

Db 916 KP-----ATSSPGTSAPSAATAAVIQNVTGQ--NIQKVAITGOLGVKPGTGN 961

Qy 300 -----LDFALE-LEFENLTPGNTNTRVSRYSSTARHLRRLRGADGTAELTTAAT--RFM 350

Db 962 SIPLTATNFRIGQKDVLRPLPSSITT-----DAKQIVLRITPDMMATLAKSQVTTVKLT 1016

Qy 351 KDLYFTSTNGVBRIGRIALTFLFNADTLGLGIPTELISAGGOLFYSRPPVVSANGPPTV 410

Db 1017 QDLFGTGGN--TTKGISATLHTSN-----PVAADSPAKASSASAPSTPTGTTVV 1067

Qy 411 KLYTSVENAQQDK-----GTAIPIHDIDLGSERVVIQDYNQHQEDRPTTSPAPSRP 461

Db 1068 KVTPLDKPTTASSAFRLMPALGVSV-----ADQKSTVASSAKP 1109

Qy 462 FSVLRANDVLMWL-----SLTAAEYDQ--STYGSSTGPVYVSDVTL-----VN 502

Db 1110 AATIRIVQGLGVMPKPAQGITVATHAKQGNASVSGSTVHTS-AVSLPSMNAAVSKTVA 1168

Qy 503 VATGAQAVARSLDW-----TKVLIDGRPLSTIQOYSKTFV 538


```
Db 1169 VASGAASSTPISGTGAPTRQVFPVSTVSTQAGKLPTRITV-----PLSVISQ----- 1218
QY 539 LPLRGK-----LSFWAGT--TKAGYPYNYNTASDQLLVE 572
Db 1219 -PMKGSVVTAPIKGNLGNLSGLGRNIIITMPAGTKLIAGNKPVSFLTAAQLOOLOQ 1277
QY 573 NAAGHRVAISTYVTS-LGAGPVSISAVAV 600
Db 1278 QCATQVRIQTPASHLQQGTAGSSSKAV 1306

RESULT 8
hemagglutinin-like secreted protein XF0889 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82750
R:anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; PMID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82750
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3282 <SIM>
A:Cross-references: GB:AE003928; GB:AE003849; NID:g9105798; PIDN:AAF93699.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, C.; El-Dorry, H.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H
as-Neto, E.; Docena, C.; M.R.-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sancelli, R.V.; Sawasak
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0889

Query Match 4.0%; Score 135.5; DB 2; Length 3282;
Best Local Similarity 21.9%; Pred. No. 7.4;
Matches 135; Conservative 64; Mismatches 216; Indels 201; Gaps 27;

QY 97 ASRRRPTTAGAAPTAVAPADTP-----PVP-----DVDSRGAIL 132
Db 429 ASPADPT-----PVASVPPTDTPSTGTMPISHPQAPVLAAGRLLTLENDIDNRGGHI 483
QY 133 RQYNLSTSPITSSVATGNLVYAAPLSPL-LPLQDGTNTH---IMATEA-----SN 181
Db 484 TAGGAIDAILTLNDRGT-----AALNRLTQOQLDNOHQILTEATHTHTLNN 537
QY 182 YQYRVARATIRYPLVNAVGVYAISFWPQTTTTPTSDVMSITSDVRLVQPGIA 241
Db 538 AAGQLHVNGTLDLTAQHLNNAAGHLLTTG--PQSATL-TIADLLDNTSGTL-----ASA 588
QY 242 SELVIPERLHYRNGQWRSVETSGVAEEATSGVLMCIHGLSVNSYNTPTGALGLLD 301
Db 589 GSLFLTAATL-----DITDGTVQSGQPLHMDAATLTAHRTGLTSQD 630
QY 302 FALELEFRNLTPGNNTNRSYSSTAR-----HRLRRGA----- 335
Db 631 -----TLTLTGHTDLSHATTAAQI TLHTGTLTAGGHTAYGHTHQLERFLD 682
QY 336 -----DGTAEI LTTTAATPMKDL YFTSNGVGEI GRGIAL TLFNLA DTL LGLSPTEL 387
Db 683 NTGQTATNGTLDLHTAALDNTGGTLHSTAT---GPNRLDITDTLTNTAGHLLNGPTTL 739
```

```
QY 388 I----SSAGQLFYSRPVV-----SANGPTVKLYTSVENAQDQKGIAPH 429
Db 740 TTGTWNTGHLQITQIPATLHATTLNDRGILDTATGPLDRLTGTLLNN--QDNGLLTTT 797
QY 430 D-IDLGESRVVIQDYDNOH-EQDRPTSPAP-----SRPFSLRVLRANDVWL----- 473
Db 798 DALTLTAASLL-----NQHTLDAAGPAHLTLTGLDNTAGLTLQTAHTLWLSAGLTNRS 852
QY 474 -SLTAEDYDQSTYGS-----TGPV-----YVSDSVTL 500
Db 853 GTLTASQLTLDTCQHTLNDTRGRGLTTCGNLTLSGFLDNTAGLTLQTAALTTDTGAATL 912
QY 501 VNVATGAQAVASRLDWTKTLDGRPLSTIQYQSKTFVPLRQKLSFWAGTTKAGYPYN 560
Db 913 TNRDGGALLAATLALDLTATLDRNG-GTIDSQTAT-----HLHTTLDNTTAGHISS 963
QY 561 YNTTASDQLLVENAAAG 576
Db 964 SGTQLQIDGTTLNTGG 979

RESULT 9
hemagglutinin-like secreted protein XF2196 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82589
R:anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; PMID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82589
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3442 <SIM>
A:Cross-references: GB:AE004032; GB:AE003849; NID:g9107324; PIDN:AAF84995.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, C.; El-Dorry, H.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H
as-Neto, E.; Docena, C.; M.R.-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sancelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2196

Query Match 4.0%; Score 135.5; DB 2; Length 3442;
Best Local Similarity 21.9%; Pred. No. 7.9;
Matches 135; Conservative 64; Mismatches 216; Indels 201; Gaps 27;

QY 97 ASRRRPTTAGAAPTAVAPADTP-----PVP-----DVDSRGAIL 132
Db 429 ASPADPT-----PVASVPPTDTPSTGTMPISHPQAPVLAAGRLLTLENDIDNRGGHI 483
QY 133 RQYNLSTSPITSSVATGNLVYAAPLSPL-LPLQDGTNTH---IMATEA-----SN 181
Db 484 TAGGAIDAILTLNDRGT-----AALNRLTQOQLDNOHQILTEATHTHTLNN 537
QY 182 YQYRVARATIRYPLVNAVGVYAISFWPQTTTTPTSDVMSITSDVRLVQPGIA 241
Db 538 AAGQLHVNGTLDLTAQHLNNAAGHLLTTG--PQSATL-TIADLLDNTSGTL-----ASA 588
QY 242 SELVIPERLHYRNGQWRSVETSGVAEEATSGVLMCIHGLSVNSYNTPTGALGLLD 301
```

Db 589 GSLTLTAATL-----DITDGTQSGGQGPLHMDAATLTAHKGTLTSD 630
QY 302 FALELEFRNLTPGNTNTRVRSYSSSTAR-----HRLRGA----- 335
Db 631 -----TLTGTHTDLSHATTTAAQIILHTGDLTTAGGHLTAAGTHTLQLEARTLD 682
QY 336 -----DGTABLTAAATRFMKDLYFTSTNGVGVIGRGIALTLNADTLTGGIPLTEL 387
Db 683 NTGGTIATNGTLDLHTAALDNTGGTLHSTAT---GPNRLDITDTLTNTAGHLLNGPTTL 739
QY 388 I---SSAGGQLFYSRPVV-----SANGPTVKLYTSVENAQDQKGLAIPH 429
Db 740 TTGTWNTTNGHLQITGPAHLTAHLLDNRGILHTATGDLRLITGTLNN--QNGLLT 797
QY 430 D-IDLGESRVVIQDYNQH-EQDRPTSPAP-----SRPFSVLRANDVWL 473
Db 798 DALTLTAASLL-----NQGTLDAAAGPAHLTLTGLLDNTAGLLQTAHLLWLTSGLTNRS 852
QY 474 -SLTAAEYDQSTYSS-----TGPV-----YVSDSVTL 500
Db 853 GTLTASQLTDLTQGHLDNTRGRGLTTGNTGLHSGPLDNTAGLLQTAALTTDTGAATL 912
QY 501 VNVATGAQAVASLDWTKVTLDCRPLSTIQYQSKTFVFLRGLKLSWEAGTTKAGYPYN 560
Db 913 TNRDGGALLAATLDTTATLDNRG-GTIDSQTAT-----HLHTTLDNTAGHISS 963
QY 561 YNTTASDQLLVENAAAG 576
Db 964 SGTLQIDGTLTWTGG 979

RESULT 10

B82519 hemagglutinin-like secreted protein XP2775 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; PMID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below
A;Accession: B82519
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-3455 <SIM>
A;Cross-references: GB:AE004082; GB:AE003849; NID:G9108003; PIDN:AAF85560.1; GSPDB:GN001
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
Rodrigues, V.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2775

Query Match 4.0%; Score 135.5; DB 2; Length 3455;

Best Local Similarity 21.9%; Pred. No. 8;
Matches 135; Conservative 64; Mismatches 216; Indels 201; Gaps 27;

QY 97 ASRRPPTTAGAAPLTAAPAHDTP-----PVP-----DVDSRGAIL 132
Db 429 ASPADPT-----FVASVPPATDTPSTGTMPPIPSHPQAPVLAAGRLTLENDIDNRGGHI 483

QY 133 RQRYNLSTSPLTSSVATGNLVLYAAPLSPL-LPQDGTNTH---IMATEA-----SN 181
Db 484 TAGGADAILTTLDNRDGT-----AALNRLTQGLQDQHGILLTATEATHHTLNN 537
QY 182 YQYRVARATIRYRLPVNAVGGYAISSISFWPQTPTTPTSDMNSITSTDRILLVOPGIA 241
Db 538 AAGQLHVNGTLDLTAQHLLNNSGHLITG--FQSATL-TIADLLDNTSGTL-----ASA 588
QY 242 SELVTPSERLHYRNQWRSVETSGVABEATSGLYMLCTHGSIVNSYNTPTTGALGLLD 301
Db 589 GSLTLTAATL-----DITDGTQSGGQGPLHMDAATLTAHKGTLTSD 630
QY 302 FALELEFRNLTPGNTNTRVRSYSSSTAR-----HRLRGA----- 335
Db 631 -----TLTGTHTDLSHATTTAAQIILHTGDLTTAGGHLTAAGTHTLQLEARTLD 682
QY 336 -----DGTABLTAAATRFMKDLYFTSTNGVGVIGRGIALTLNADTLTGGIPLTEL 387
Db 683 NTGGTIATNGTLDLHTAALDNTGGTLHSTAT---GPNRLDITDTLTNTAGHLLNGPTTL 739
QY 388 I---SSAGGQLFYSRPVV-----SANGPTVKLYTSVENAQDQKGLAIPH 429
Db 740 TTGTWNTTNGHLQITGPAHLTAHLLDNRGILHTATGDLRLITGTLNN--QNGLLT 797
QY 430 D-IDLGESRVVIQDYNQH-EQDRPTSPAP-----SRPFSVLRANDVWL 473
Db 798 DALTLTAASLL-----NQGTLDAAAGPAHLTLTGLLDNTAGLLQTAHLLWLTSGLTNRS 852
QY 474 -SLTAAEYDQSTYSS-----TGPV-----YVSDSVTL 500
Db 853 GTLTASQLTDLTQGHLDNTRGRGLTTGNTGLHSGPLDNTAGLLQTAALTTDTGAATL 912
QY 501 VNVATGAQAVASLDWTKVTLDCRPLSTIQYQSKTFVFLRGLKLSWEAGTTKAGYPYN 560
Db 913 TNRDGGALLAATLDTTATLDNRG-GTIDSQTAT-----HLHTTLDNTAGHISS 963
QY 561 YNTTASDQLLVENAAAG 576
Db 964 SGTLQIDGTLTWTGG 979

RESULT 11

T21460

hypothetical protein ZK945.10 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T21460

R;Wilkinson, J.

submitted to the EMBL Data Library, March 1995

A;Reference number: Z19425

A;Accession: T21460

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-796 <WTL>

A;Cross-references: EMBL:Z48582; PIDN:CAA8469.1; GSPDB:GN00020; CESP:ZK945.10

A;Experimental source: clone F27E5

C;Genetics:

A;Gene: CESP:ZK945.10

A;Map position: 2

A;Introns: 62/3; 124/3; 169/1; 216/1; 262/3; 371/1; 581/3; 608/1

Query Match 3.9%; Score 133; DB 2; Length 796;

Best Local Similarity 20.6%; Pred. No. 1.4;

Matches 121; Conservative 61; Mismatches 228; Indels 176; Gaps 24;

QY 63 APDVTAAGAGPRVQPARPLGSAMRQQAORPAVARRRPTTAGAAPLTAAPAHDTPPV 122
Db 242 AKQFSMRTSGSPTLRRMKRDAGDNTCDYTIESTSTSTTPTTT-----TVTSTVTSTTV 296

QY 123 PDVDSRGAILLRQYNLTSPLTSSVATGNLVLYAAPLSPLPQDGTNTHIMATEAGNY 182
Db 297 P-----TSTSTVTVTAMSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 338

QY 109 PUTAVAPAHDTTPVPVDSRGAILRR-----QYNLSTSP-----TSS 146
Db 28 PTTSSVTSADPTTSADPTTSVQTTSAGPSNNIGNSTLANSTTFAVSTSIDPTSSD 87
QY 147 VATGTNLVLYAAPLSPLLPQD---CTNTHIMATEASNYAQYVARATIRYRPLVNAV 203
Db 88 VITSTVQTTSIEPTTSLVSSNDITTSNLINISVVIISTDSTSLIESTTTVGHASSVA 147
QY 204 G-YAISISFPWPTTPTSDMNSITSTDRILVQPGIAS---ELVIPSERLHYRQGW 259
Db 148 GMYRTSSDEVITSTDPST-----SSDVASADPTSSSAVTTLVDPPTSV----- 192
QY 260 SVETSGVAEEATSGVLMLCIHGSLVNSYNTPTGALGLLDPALEFRNLTPGNVTR 319
Db 193 -VISTSVDOTSSSDVATSDPTTSVISTSSADPTTSA-----DSTTSTVQTTS 239
QY 320 VSRYSSTARHLRRGADGTAELTTAATRFMKDLYFTSTNGVGEIGKGIALTFLNLADTL 379
Db 240 VDTSSVVSAPVDPASSVVSLSSTSSVPTSSSTVITISANGSATLA-----AQIT 289
QY 380 LGLLPTTELISAGGQLFYGRP-----VVSANGPTVKLYTSVE-NAQODKGIAPHD 430
Db 290 SIDPVSSIVSSGATTIISASIDPASSVVSSTSSSEPTSFIVSSTSVSTRPSGPTTSTD 349
QY 431 IDIGESRVVIQDYDQHEQDRPTSPAPRPFVLRANDVLMLSLTAAYDQSTYGSSTG 490
Db 350 LATFSDTIILRVTTSTSTQDTQTVSSSLTDMVSSSTGSDLSVSIQRSQVDPSTFAVNS 409
QY 491 PVTVSDSVTLVNVATGAQAVARSLDWTK-----VTLDGRPLSTIQYKSTFFV 538
Db 410 PVYPTASTG---SSTGTPIASLSLSRQOGISATSSSVITLTPVDSASSSSRSATSII 467
QY 539 LP-----LRGKLSFWEA-GTTKAGYP-----YNYNTASDQLIV-ENAAHRVAIS 582
Db 468 KPNWPSSNDKTKQSSVVVDAPQSTKSSYPSTISADPTTLASENGLVGSSSAHPITLD 527
QY 583 -TYTTSLGAGPVSISAVAVLAPHAL 607
Db 528 RTYASAHASVTDIVSRVTDSTRHTTL 553

RESULT 14

S59310

probable membrane protein YMR317w - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein YM9924.09

C;Species: Saccharomyces cerevisiae

C;Date: 29-Nov-1995 #sequence_revision 23-Feb-1996 #text_change 19-Apr-2002

C;Accession: S59310

R;Churche, C.M.

submitted to the EMBL Data Library, September 1995

A;Reference number: S59302

A;Accession: S59310

A;Molecule type: DNA

A;Residues: 1-1104 <CHU>

A;Cross-references: EMBL:Z54141; GSPDB:GN00013; MIPS:YMR317w

A;Experimental source: strain AB972

C;Genetics:

A;Gene: MIPS:YMR317w

A;Cross-references: SGD:S0004936

A;Map position: 13R

Query Match 3.8%; Score 130.5; DB 2; Length 1104;
Best Local Similarity 18.7%; Pred. No. 3.2; Indels 121; Gaps 22;
Matches 113; Conservative 109; Mismatches 262

QY 63 APDVTAAAGAPRVQRPARPLGSARWDOAQRPAVASRRPTTAGAAPLTAVAPA-----H 117

Db 148 SPQTSSSNGG-----SSSEPLGK-----SSVLETTASSDDTAVTSFTTTL 191

QY 118 DTPVPVDSRGAIL-----RQYNI.LST-----SPLTSSVA-----TGT 151

Db 192 DVSSSPKISSGSAVTSVGTTSASKEYFSSSTSDVSSLLSTSSPASSTISLTPFSST 251

QY 152 NLVLVYAAPLSPLLPQDGTNTHIMATEASNYAQYVARATIRYRPLVNAVGGYAIISI-- 209
Db 252 ILUSTSSPVSSSEAP--SATSSSVSEASSTSSSVSEAPLATSSVVSSEAPSTSSVVS 309
QY 210 SEWPQTTTTPTSDMNSITSTDRILVQPGIASLSEVIPSE-----RLH 252
Db 310 SEAPSTSSSVSEISSTSSSVS--SEAPLATSSVVSSEAPSTSSSVSSEAPSTSSS 367
QY 253 YNQGWRSVETSGVAEEATSGVLMLCIHGSLVNSYNTPTGALGLLDPALEFRNL 312
Db 368 VSSEAPSTSSSVSEISSTKSVMSSEVSSATSSVLSSEAPSAISL--ASSRLPSSKN 425
QY 313 PGNTNTRVRSYSTARHLRRRGADGTAELTTAATRFMKDLYFTSTNGVGEIGRGIA 369
Db 426 TSVTSLVATEASSVTSLSRPSSETLASNIIESSLSTGVNSVSTVTTSAASLTGSKVS 485
QY 370 LTLFNLADTLGGLPTELISAGGQLFYSRPVVSANGPTVKLYTSVENAQDQKGAIP- 428
Db 486 SSSNRMATKSTSTSSDL--SKSSVIFGNSSTVTTSPSASISLTAS-----PLPS 533
QY 429 --HDIDLGESRVVIQDYDQHEQDRPTSPAPRPFVLRANDVLMLSLTAAYDQSTY 486
Db 534 VMSDITSSSEASSISNL-----ASSAPSDDNNSTIAS-----ASLLVTKIKSVVS 579
QY 487 SSTGPVVVSDSVTLNVNATGA-----QAVARSLDWTKVT-LDGRPLSTIQYK-TFFV 538
Db 580 SVSSITSETNESNLATSSSTLLSNKATARSLSSTNATSASNVPTGTFSSWSHTSVI 639
QY 539 LPARGKLSFWEAGTTKAGYPYNYNTASDQLIVENAAHRVAISVTTTSLGAGPVSISAV 598
Db 640 TP-----GFSTSGASLAIN-STVVSSSLAGYSFSTPSSPTTSTLTSEAPSTVSSM 690
QY 599 AVLAP 603
Db 691 TTSAP 695

RESULT 15

S57180

probable membrane protein YJR151c - yeast (Saccharomyces cerevisiae)

N;Alternate names: protein J2223; serine/threonine-rich protein YJR151c

C;Species: Saccharomyces cerevisiae

C;Date: 23-Aug-1995 #sequence_revision 08-Sep-1995 #text_change 19-Apr-2002

C;Accession: S57180

R;Scarcez, T.

submitted to the Protein Sequence Database, September 1995

A;Reference number: S57169

A;Accession: S57180

A;Molecule type: DNA

A;Residues: 1-1161 <SCA>

A;Cross-references: EMBL:Z49651; NID:g1015902; PID:g1015903; GSPDB:GN00010; MIPS:YJR151c

C;Genetics:

A;Gene: SGD:YJR151c

A;Cross-references: SGD:S0003912

A;Map position: 10R

C;Keywords: transmembrane protein

Query Match 3.8%; Score 130.5; DB 2; Length 1161;
Best Local Similarity 17.2%; Pred. No. 3.4; Indels 133; Gaps 21;
Matches 104; Conservative 95; Mismatches 274

QY 57 HPTNPFAPDVTAAAG-----AGPRVRQPARPLGSARWDOAQRPAVASRRR--- 101

Db 59 HKETYPSEIAAAVFDYGDFTTRLTGIGSDEVTRMI--TGVPMYSTRKPKAISALSADG 116

QY 102 -----PTTAGAAPLTAVAPAHDTTPVPVDSRGAILRQYNI.LSTSSVA-----TGTN 152

Db 117 IYTAIPTSTSTTTTKSSTSTPTTITSTTSTTPTTSTTPTTSTTPTTSTTPTTSTT 176

QY 153 LVLYAAPLSPLLPQDGTNTHIMATEASNYAQYVARATIRYRPLVNAVGGYAIISFW 212

Db 177 STPTTSTTSTTPTTSTTST-----TPTTSTTSTTPTTSTTSTTPTTSTTSTT 224

Blank Sheet


```
Db 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRLVFNAGGYAISIFPQTTTTPTSDVMNSITSDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRLVFNAGGYAISIFPQTTTTPTSDVMNSITSDVRLVQPGI 240
QY 241 ASELVIPSRLHYRNQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 300
Db 241 ASELVIPSRLHYRNQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 300
QY 301 DPALELFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEELTTTAATRFMKDLYFTSTNG 360
Db 301 DPALELFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEELTTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEQRPTSPAPSRPFVLRANDVWLJSLTAAEY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEQRPTSPAPSRPFVLRANDVWLJSLTAAEY 480
QY 481 DOSTYSSGTPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIQQYSKTFFVLP 540
Db 481 DOSTYSSGTPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIQQYSKTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHGRAHVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHGRAHVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDLDYPARAHTPDDFCPCRPGLGQCAFOSTVAELQRLKMKVGKTRREL 660
Db 601 LAPHSALELLEDLDYPARAHTPDDFCPCRPGLGQCAFOSTVAELQRLKMKVGKTRREL 660
```

RESULT 2

```
VST2_HEVPA
ID VST2_HEVPA STANDARD; PRT; 660 AA.
AC P33426;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=33774;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M80581; AAA45727.1; ..
CC InterPro; IPR004261; SP2.
CC DR InterPro; IPR008975; Viral_cap_coat.
CC DR Pfam; PF03014; SP2; 1.
CC
CC KW Signal.
CC FT SIGNAL
CC CHAIN 23 660 BY SIMILARITY.
CC
```

```
SQ SEQUENCE 660 AA; 70980 MW; 8085BC53CEB46FD3 CRC64;
Query Match 99.5%; Score 3391; DB 1; Length 660;
Best Local Similarity 99.4%; Pred. No. 2.7e-207;
Matches 656; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGGFWMGDRVDSQFALPYIHTPN 60
Db 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGGFWMGDRVDSQFALPYIHTPN 60
QY 61 PFAPDVTAAGAGPRVRQPARPLGSARWQAQAPAVASRRRPTTAGAAPTAVAPADTP 120
Db 61 PFAPDVTAAGAGPRVRQPARPLGSARWQAQAPAVASRRRPTTAGAAPTAVAPADTP 120
QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLLPLQDGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRLVFNAGGYAISIFPQTTTTPTSDVMNSITSDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRLVFNAGGYAISIFPQTTTTPTSDVMNSITSDVRLVQPGI 240
QY 241 ASELVIPSRLHYRNQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 300
Db 241 ASELVIPSRLHYRNQWRSVETSGVAEEATSGLVMLCIHGSPVNSYNTPTYGALGLL 300
QY 301 DPALELFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEELTTTAATRFMKDLYFTSTNG 360
Db 301 DPALELFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEELTTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEQRPTSPAPSRPFVLRANDVWLJSLTAAEY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEQRPTSPAPSRPFVLRANDVWLJSLTAAEY 480
QY 481 DOSTYSSGTPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIQQYSKTFFVLP 540
Db 481 DOSTYSSGTPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIQQYSKTFFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHGRAHVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHGRAHVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDLDYPARAHTPDDFCPCRPGLGQCAFOSTVAELQRLKMKVGKTRREL 660
Db 601 LAPHSALELLEDLDYPARAHTPDDFCPCRPGLGQCAFOSTVAELQRLKMKVGKTRREL 660
```

RESULT 3

```
VST2_HEVMY
ID VST2_HEVMY STANDARD; PRT; 660 AA.
AC Q04611;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227573; PubMed=8470371;
RA Aye T.T., Uchiida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
RL Virus Genes 7:95-109(1993).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC
```



```
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10330; BAA01174.1; -
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
KW SIGNAL.
FT CHAIN.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 660 AA; 70998 MW; 3A82A4EA25C6253 CRC64;

Query Match
Best Local Similarity 98.8%; Score 3367; DB 1; Length 660;
Matches 653; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRPRDILLILLMLPMLPAPPQSGRRRSGSGGFGWDRVDSQPPFAIPYIHPTN 60
DB 1 MRPRDILLILLMLPMLPAPPQSGRRRSGSGGFGWDRVDSQPPFAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHPTP 120
DB 61 PFAPDVTAAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHPTP 120
QY 121 PVPDVSRRGAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
DB 121 PVPDVSRAAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYRPLVPNAVGGVAISISFWPQTTTPTSDVMNSITSTDVRIIVQPGI 240
DB 181 NYAQVRVARATIRYRPLVPNAVGGVAISISFWPQTTTPTSDVMNSITSTDVRIIVQPGI 240
QY 241 ASELVIPSERLHYRNQGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGL 300
DB 241 ASELVIPSERLHYRNQGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGL 300
QY 301 DFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
DB 421 QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
QY 481 DQSTYSGSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGDRPLSTIQYKTFVFLP 540
DB 481 DQSTYSGSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGDRPLSTIQYKTFVFLP 540
QY 541 LRKLSFWEAGTTKAGYPYNTNTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVV 600
DB 541 LRKLSFWEAGTTKAGYPYNTNTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVV 600
QY 601 LAPHASALLEDTLDYPACAHFTDDFCPCPLGIGQCAFOSVTAELORLKMVKGTREL 660
DB 601 LAPHASALLEDTLDYPACAHFTDDFCPCPLGIGQCAFOSVTAELORLKMVKGTREL 660

RESULT 4
VST2 HEVME
ID VST2 HEVME STANDARD; PRT; 659 AA.
AC Q03500;
DT 01-OCT-1993 (Rel. 27, Created)
```

```
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor.
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31768;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93079857; PubMed=1448913;
RA Huang C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E.,
RA Bradley D.W., Tam A.W., Reyes G.R.;
RT "Molecular cloning and sequencing of the Mexico isolate of hepatitis
RT E virus (HEV)".
RL Virology 191:550-558(1992).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M74506; AAA45732.1; -
DR PIR: B44212; B44212.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
KW SIGNAL.
FT CHAIN.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 659 AA; 70640 MW; CF75E75EFD8FBE2C CRC64;

Query Match
Best Local Similarity 93.7%; Score 3194.5; DB 1; Length 659;
Matches 614; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

QY 1 MRPRPILLILLMLPMLPAPPQSGRRRSGSGGFGWDRVDSQPPFAIPYIHPTN 60
DB 1 MRPRPILLILLMLPMLPAPPQSGRRRSGSGGFGWDRVDSQPPFAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHPTP 120
DB 61 PFAPDVTAAAGAGPRVROPAPPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPAHPTP 120
QY 121 PVPDVSRRGAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
DB 121 PVPDVSRAAILRRQYNLSTPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYRPLVPNAVGGVAISISFWPQTTTPTSDVMNSITSTDVRIIVQPGI 240
DB 181 NYAQVRVARATIRYRPLVPNAVGGVAISISFWPQTTTPTSDVMNSITSTDVRIIVQPGI 240
QY 241 ASELVIPSERLHYRNQGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGL 300
DB 241 ASELVIPSERLHYRNQGWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGL 300
QY 301 DFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALEFRNLTPGNTNTRVSRYSSTARHS -ARGADGTAEITTTAATRFMKDLFTGLNG 359
QY 361 VGEIGRGIALTLFNLADTLGLLPTLISAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLNADTLGLLPTLISAGGQLFYSRPVVSANGPTVKLYTSVENAQ 419
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
DB 420 QDKGVAIPHDIDLGRSVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVWLSLTAAEY 479
```

QY 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSMDWKVTLTDCRPLSTIQOYKSKTFVLP 540
DB 480 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSMDWKVTLTDCRPLSTIQOYKSKTFVLP 539
QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVNNAAGHRVAISTYTTSLGAGPVSISAVAV 600
DB 540 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVNNAAGHRVAISTYTTSLGAGPVSISAAV 599
QY 601 LAPHSALALLEDTLDYPARAHTFDDFCPCRCPLGLQGCFAFQSTVABLRKMKVKGKREL 660
DB 600 LAPRSALALLEDTLDYPARAHTFDDFCPCRCALGLOGCAFQSTVABLRKMKVKGKREL 659
RESULT 5
VST2 HEVRH STANDARD; PRT; 485 AA.
AC Q00270;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 01-OCT-1993 (Rel. 27, Last annotation update)
DE Structural protein 2 (Fragment).
OS Hepatitis E virus (isolate Rhesus) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31766;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K.,
RA Win K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90274; BAA20910.1; --
CC InterPro; IPR004261; SP2.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF03014; SP2; 1.
CC FT NON_TER 1 485
CC FT NON_TER 485 485
CC FT SEQUENCE 485 AA; 52317 MW; 5A4FD9273AC74F94 CRC64;
Query Match 71.8%; Score 2446; DB 1; Length 485;
Best Local Similarity 99.0%; Pred. No. 1.2e-147;
Matches 480; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 133 RQYNLSTSPSTSSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEASNAQYVARATI 192
DB 1 RQYNLSTSPSTSSVATGNTLVLYAAPLSPLLPQDGTNTHIMATEASNAQYVARATI 60
QY 193 RYRPLVNAVGGYIAISISFWPQTPTTPTSDVMSNITSDVRLVQPGIASLVIPSERLH 252
DB 61 RYRPLVNAVGGYIAISISFWPQTPTTPTSDVMSNITSDVRLVQPGIASLVIPSERLH 120
QY 253 YRNOGWSRVETSGVAEBEATSGLVNMLCIHGLSVNSYNTPTGTGALLDFALEFRNL 312
DB 121 YRNOGWSRVETSGVAEBEATSGLVNMLCIHGLSVNSYNTPTGTGALLDFALEFRNL 180
QY 313 PGNTNTRSVRSYSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNGVGBIGGIALTL 372
DB 181 PGNTNTRSVRSYSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNGVGBIGGIALTL 240
QY 373 FNLADTLTGGLPTELLISSAGGQLFYSRPWSANGEPYVKLYTSVVENAQDQKGIAPHDID 432
DB 241 FNLADTLTGGLPTELLISSAGGQLFYSRPWSANGEPYVKLYTSVVENAQDQKGIAPHDID 300

QY 433 LGSERVIQDYNQHEQDRPTSPAPSRPFSVLRANDVLMLSITAAEYDQSTVGSSTGPV 492
DB 301 LGSERVIQDYNQHEQDRPTSPAPSRPFSVLRANDVLMLSITAAEYDQSTVGSSTAPV 360
QY 493 YVSDSVTLNVNATGAQAVARSMDWKVTLTDCRPLSTIQOYKSKTFVLPRLGKLSFWEAGT 552
DB 361 YVSDSVTLNVNATGAQAVARSMDWKVTLTDCRPLSTIQOYKSKTFVLPRLGKLSFWEAGT 420
QY 553 TKAGYPYNYNTTASDQLLVNNAAGHRVAISTYTTSLGAGPVSISAVAVLAPHSALELLED 612
DB 421 TKAGYPYNYNTTASDQLLVNNAAGHRVAISTYTTSLGAGPVSISAVAVLAPHSALELLED 480
QY 613 TLDYP 617
DB 481 TLDYP 485
RESULT 6
EGT2_YEAST STANDARD; PRT; 1041 AA.
AC P42835;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE EGT2 protein precursor (Early G1 transcript 2).
GN EGT2 OR YNL327W OR N0320.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9288C / FY1676;
RX MEDLINE=95373280; PubMed=7645347;
RA Maftahi M., Nicaud J.-M., Levesque H., Gaillardin C.;
RT "Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV
RT identifies the RPD3, PAS8 and KRE1 loci, five new open reading
RT frames.";
RL Yeast 11:567-572(1995).
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=96251274; PubMed=8668141;
RA Kovacech B., Nasmith K., Schuster T.;
RT "EGT2 gene transcription is induced predominantly by Swi5 in early
RT G1.";
RT Mol. Cell. Biol. 16:3264-3274(1996).
CC -|- FUNCTION: Seems to be involved in the correct timing of cell
CC separation after cytokinesis, as separation of mutant daughter
CC cells is delayed. Could either be an enzyme necessary for glucans-
CC degradation of the cell wall at the neck region between mother and
CC daughter cells or a regulatory protein controlling this metabolic
CC step. Exclusively expressed between the end of mitosis and early
CC G1; inactivated before cells pass start.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z45259; CAA86371.1; --
CC EMBL; Z71603; CAA96259.1; --
CC PIR; S55862; S55862.
CC GermOnline; 143333; --
CC SGD; S0005271; EGT2.
CC GO; GO:0009277; C:cell wall (sensu Fungi); IDA.
CC GO; GO:0000910; P:cytokinesis; IMP.
CC GO; GO:0000910; P:cytokinesis; IMP.
KW Glycoprotein; Repeat; Signal; Cell cycle; Cell division; Mitosis.
FT SIGNAL 1 20
FT CHAIN 21 1041 EGT2 PROTEIN.


```
FT CARBOHYD 4965 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 4987 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5037 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5052 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5156 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5166 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5427 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5467 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5506 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5507 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5543 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5553 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5604 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5618 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT CARBOHYD 5662 N-LINKED (GLCNAC. . .) (POTENTIAL. . .)
FT VARIANT 5137 T -> S (in dbSNP:2672788).
/FTId=VAR_014123.
FT 34 G -> E (IN REF. 2).
FT 95 FPGLCN -> LPCLCK (IN REF. 2).
FT 104 S -> C (IN REF. 2).
FT 142 E -> K (IN REF. 1).
FT 225 R -> S (IN REF. 2).
FT 331 PL -> T (IN REF. 2).
FT 337 E -> N (IN REF. 2).
FT 356 E -> K (IN REF. 2).
FT 362 G -> R (IN REF. 2).
FT 369 MISSING (IN REF. 2 AND 3).
FT 374 D -> N (IN REF. 2).
FT 394 RT -> TR (IN REF. 2).
FT 468 RK -> GR (IN REF. 2).
FT 512 L -> P (IN REF. 2).
FT 585 GAA -> AH (IN REF. 3).
FT 601 A -> S (IN REF. 3).
FT 628 DP -> RS (IN REF. 2).
FT 633 F -> L (IN REF. 2).
FT 676 A -> P (IN REF. 3).

Query Match 4.0%; Score 137; DB 1; Length 5703;
Best Local Similarity 20.1%; Pred. No. 5.1;
Matches 143; Conservative 72; Mismatches 257; Indels 240; Gaps 33;

QY 52 AIPYIH--PTNFPADVTAAGAPRVQPARPLGSAWRDQAPAVASRRRTTGAAP 109
DB 3192 ALPALRSTATTPTATSTVATL-----PSSSLGTAW-----TRLSQTTTPTAT 3232
QY 110 LTAVAHADTPPVDPVDVDSRAILRRQNLSTPLTSVATGTVLVAAPLSPLLPQDG 169
DB 3233 MSTATPS-STPE-----TWHTSTVLTTTATTTRGVSATPSS-----TPG 3271
QY 170 T--NTHIMATEASNYAQYRVARATIRYRPLVPNAVGGYALISGFWPOTTITPT----- 220
DB 3272 TAHTTKVPITTTTGFT-----ATPSSSPGALTTPPWISTTTTPTTRGSTVT 3318
QY 221 --SVDNNSITSDVRLVQPGIASELVIPSERLHYRNQGWRSVETSGVAEEATSGLVML 278
DB 3319 PSSIPGTHTATVLTITTTTAVATGSMATPSS-----STQTSPTPSLTTATTIT 3368
QY 279 CIHGLSVN-----SVTNPYTGALG-----LLDPALELEF 308
DB 3369 AT--GSTNPSSTGTPPIPVLTATTATPAATSTVTP--SALGTHTPPVPNATTATHG 3426
QY 309 RNLTPGNTRVRYSGSTARHLRGADGTAEIT-----TTAAT----- 347
DB 3427 RSLPPSPHTVPTAWTSAT-----SGILGTHITEPTSGTSHTPAATTTGTPSTPALSS 3481
QY 348 -----RPMKDLTYSTNGVEIGRGIALTLFNL-----ADTLGLGLPTE-----LISAG 392
DB 3482 PHFSSRTTESPPPGTTPPGHT-RGTSRTTATATTPSKRTSTLLSPSPTSAPITTVVTTG 3540
QY 393 GO-----LFYSRPVWSANGPTEVKLYTSVE-----NAQDQGI----- 425
DB 3541 CEFQCAWSEWLDYSWMPGPGSG-GDFDTYSNIRAAGGAVCEQPLGLECRQAQGVPLRE 3599
QY 426 ---AIPHDIDLG-----ESRVVIQDYNQHQDRP-----TPSPAP 458
```


RA Kim K., Bajszar G., Lee S.Y., Knudsen P., Mattoon J.R.;
 RT "Cloning of a new allelic variant of a Saccharomyces diastaticus
 RT glucosylase gene and its introduction into industrial yeasts."; Appl. Biochem. Biotechnol. 44:161-185(1994).
 RL [2]
 RN
 RP
 RX
 RX MEDLINE=91276266; PubMed=2055484;
 RA Lambrechts M.G., Pretorius I.S., Sollitt P., Marmur J.;
 RT "Primary structure and regulation of a glucosylase-encoding gene
 RT (STA2) in Saccharomyces diastaticus."; Gene 100:95-103(1991).
 RL
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-
 CC glucose residues successively from non-reducing ends of the chains
 CC with release of beta-D-glucose.
 CC
 CC -1- SIMILARITY: Belongs to family 15 of glycosyl hydrolases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M60650; AAA35107.1; -;
 DR EMBL; M90490; AAA20560.1; -;
 DR PIR; J0474; J0474.
 DR HSP; P08017; IAYX.
 DR InterPro; IPR00165; Glyco hydro 15.
 DR InterPro; IPR008928; Glyco trans 6hp.
 DR Pfam; PF00723; Glyco hydro 15; 1.
 DR PRINTS; PR00736; GLHYDRLASE15.
 DR PROSITE; PS00820; GLUCOAMYLASE; 1.
 DR Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KW Signal; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 768 GLUCOAMYLASE S2.
 FT DOMAIN 22 348 SER/THR-RICH.
 FT DOMAIN 349 692 H SUBUNIT.
 FT DOMAIN 693 768 Y SUBUNIT.
 FT BINDING 456 456
 FT ACT SITE 519 519 CATALYTIC BASE (BY SIMILARITY).
 FT ACT SITE 522 522 GENERAL ACID CATALYST (BY SIMILARITY).
 FT ACT SITE 523 523 INTERACT WITH SUBSTRATES (BY SIMILARITY).
 FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 415 415 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 424 424 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 742 742 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 164 164 MISSING (IN REF. 1).
 FT CONFLICT 624 624 D -> N (IN REF. 1).
 SQ SEQUENCE 768 AA; 82586 MW; 3FAC172C129A0C6F CRC64;
 Query Match 3.8%; Score 130.5; DB 1; Length 768;
 Best Local Similarity 20.8%; Pred. No. 0.86;
 Matches 130; Conservative 70; Mismatches 231; Indels 193; Gaps 29;
 QY 4 RPILL-----LLMLPMLPAPPQPSRRRRGSGGSGGFGWDRVDSOPFALPYHPT 59
 DB 3 RPFLLAYLVLSLNFSLGPFALVP-----RGSSS-----SNITSSGSPSTPFSSAT 50
 QY 60 NPAPDVTAAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAALTAAPAHDT 119
 DB 51 ESRSTGTVT-----FSSSKYPSKSTSTSVSTTTTIVTTTSTTSTT 98
 QY 120 PPVPDVSRRGAILRRQYNLTSP--LTSVATGTN---LVLAAPLSPLPLQDGTNTHI 174

DB 99 TITTTVCSGTNSAGSTTSCSPKTIITTVPCSTSTSETASTTSTPTTPTTVVSTTV 158
 QY 175 MATEASNYAQYVARATIRYRPLVNAVGYAISL-----SWPQT-TTTPTSVD 223
 DB 159 VTTEYASTSTKQGGIITTF--VTKNPIPTLTATPTSSVTVTNFTTTTITTVCGSTG 216
 QY 224 MNSITSDVRIILVQPGIASLVLPSRLHYRNQGRSVETSGVAEEBATSGLVMLCIHS 283
 DB 217 TNSAGET-----TSGCSPKTVTT--TVPCTGT 242
 QY 284 LVNSYT---NTPYTGALGLLDFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTA 340
 DB 243 --GEYITEATAPVTAV-----TTTVTTESST-----GTNSAGK 275
 QY 341 LTTAATRRMKDLYFTSTNGVGEIGRGIALTLNLAADTLGLLPTELISAGQQLFYSRP 400
 DB 276 TTTSTYTTKSPVTTY-----VDFGKGI-----LDQSCGG---VFSNNGSSQVQLRD 318
 QY 401 VVSANGEPTVKLYTS-----VENAQODKGIAIPHDID-LGESRVVTQDYDNQHEQD 450
 DB 319 VVLWNG--TV-VYDSNGAMDSSPLEEWLQKQKVSIERIFENIGPSAVY----- 364
 QY 451 RPT-----PSPAPSRP-----FVLRANDVLMLSLTA--EYD 481
 DB 365 -PSILPGWVIASPSQTHPDYFYQWIRDSALTSIVSHSADPAIETLLQYLVNVSFHLQRT 423
 QY 482 QSTYGSSTGPVYVSDSVTLVNAVATGAQAVARSIDWTKVLDGRPLSTIQOYKTFVFLPL 541
 DB 424 NLTGAGIG--YTNDTVALGDPKNNVDNTAFTBPWGRPQNDGPALRSIA-----ILKI 474
 QY 542 RGKLSFEWAGTTRKAGYPYNYNTTA 565
 DB 475 ---IDYIKSGTDLGAKYPPQSTA 495
 RESULT 14
 DAN4 YEAST
 ID DAN4 YEAST STANDARD; PRT; 1161 AA.
 AC P47179;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cell wall protein DAN4 precursor.
 GN DAN4 OR YJR151C OR J2223.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Scareez T.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 [2]
 RP REGULATION.
 RL MEDLINE=211131168; PubMed=11160904;
 RA Cohen B.D., Sertil O., Abramova N.E., Davies K.J., Lowry C.V.;
 RT "Induction and repression of DAN1 and the family of anaerobic
 RT mannoprotein genes in Saccharomyces cerevisiae occurs through a
 RT complex array of regulatory sites."; Nucleic Acids Res. 29:799-808(2001).
 RL
 CC -1- FUNCTION: Component of the cell wall (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -1- PTM: Extensively O-glycosylated (Potential).
 CC -1- SIMILARITY: Belongs to the SRP1 / TIP1 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```
CC or send an email to license@isb-sib.ch).
CC EMBL; Z49651; CA889684.1;
DR PIR; S57180; S57180.
DR GERMOnline; 141984;
DR SGO; S0003912; DAN4.
DR GO; GO:0009277; C:cell wall (sensu Fungi); IDA.
DR Pfam; PF00660; SRP1_TIP1.
DR PROSITE; PS00724; SRP1_TIP1; 1.
KW Cell wall; Glycoprotein; Membrane; GPI-anchor; Signal; Lipoprotein.
FT SIGNAL 1 24
FT CHAIN 25 1146
FT PROPEP 1147 1161
FT LIPID 1146 1146
FT SEQUENCE 1161 AA; 118358 MW; 7954C15D69F0CA58 CRC64;
Query Match 3.8%; Score 130.5; DB 1; Length 1161;
Best Local Similarity 17.2%; Pred. No. 1.5;
Matches 104; Conservative 95; Mismatches 274; Indels 133; Gaps 21;
QY 57 HPTNPAPDVTAAG-----AGPRVQRPARPLGSHWRDQAPAVASRRR--- 101
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
59 HKTETYPSEIAAAVDYGDFTRLTGISGDEVTMRI--TGVFWYSTRLKPAISSALSXDG 116
QY 102 -----PTTAGAAPLTAVAPAHDTTPVDVDSRGAILRQYLNLSPLTSSVA---TGTN 152
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 IYALPTSTSTTTKSSSTSTPTTITSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTT 176
QY 153 LVLVAAPLPLQDQNTTHIMATEASNYAQYRVARATIRYRLVFNPAVAGYALISFW 212
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 STPTTSTSTSTPTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 224
QY 213 POTTTPSDVMNSTSTDRILVOPGIASELVPISERLHYRNQGRSVETSGVAEEBAT 272
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
225 PTTSTPTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 283
QY 273 SGLVNLCHGSLVNSVYNTPTYGALGLD--FALELEFRLNLTGNTFVRSYSTARH 330
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
284 TSTTSTAPTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 340
QY 331 LRRGADGTAEIITTAATRFMKDLYFTSNGVGEIGRGIALTLFNLAIDLGLPLTEL 390
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
341 ----VSTNAFTSATTTTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 372
QY 391 AGGQLFYRPRVVSANGBPTV--KLTVSNAQDQKGIAPHIDIDGEKRVVIQDYND 448
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
373 A-----EPTTVSEVTSSVEPTRSSQ-----VTSSAEPTTVSEPTSSVE 410
QY 449 QDRPTFPAPRPFPSVLRANDVLMLSLTAAYDQSTYGSSTGTPVYVSDSVTLVNV 508
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
411 PTRSQVTSIAEP-----TTVSEPTSSVEPTRSSQVTSIAEPPTTVSEPTSSV 457
QY 509 AVARSLDWTKVTLQGRPLSTIQYQSKTFVFLPLR--GKLSFWEAGTTKAGYPVNYNT 567
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
458 EPTRS---SQVTSIAEP--TTVSEPTSS--VEPTRSSQVTSIAEPPTTVSEPTSS 511
QY 568 QLLVNAAGHRAVISTYTTSLGAGPVISAVAVLAPHALALLEDTLTPYRAHFTDF 627
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
512 QV---TSSAEPTTVSEVTSSV--EPIRSSQVTTTEPVSSF-----GSTFSEIT 554
QY 628 PECREPL 633
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
555 SSAEPL 560
RESULT 15
PCGN HUMAN
ID PCGN HUMAN STANDARD; PRT; 1321 AA.
AC O14594; Q9UPK6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
```

10-OCT-2003 (Rel. 42, Last annotation update)
Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).
CSPG3 OR NCAN OR NEUR.
Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99013874; PubMed=9795216;
RA Prange C.K., Pennacchio L.A., Lieuallen K., Fan W., Lennon G.G.;
RT "Characterization of the human neurocan gene, CSPG3.";
RL Gene 221:199-205(1998).
RN [2]
RP SEQUENCE OF 1-990 AND 1007-1321 FROM N.A., AND VARIANT ALA-1254.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhart-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Krommiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RT "Sequence analysis of an ~1 Mb region containing the MEF2B gene in
19p12.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
development by binding to neural cell adhesion molecules (NC-CAM
and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
acid.
CC -!- TISSUE SPECIFICITY: Brain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF026547; AAC80576.1; -;
EMBL; AC003110; AAB86655.1; -;
EMBL; AC005254; AAC25581.1; -;
HSSP; P00740; 1EDM.
GeneW; HGNC:2465; CSPG3.
MIM; 600826; -;
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00059; lectin_C; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR PRODOM; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.

DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE LECTIN 1; 1.
DR PROSITE; PS00041; C_TYPE LECTIN 2; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal; Polymorphism.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 1321 NEUROCAN CORE PROTEIN.
FT DOMAIN 38 153 IG-LIKE V-TYPE.
FT DOMAIN 159 254 LINK 1.
FT DOMAIN 260 356 LINK 2.
FT DOMAIN 1008 1044 EGF-LIKE 1.
FT DOMAIN 1046 1082 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1084 1213 C-TYPE LECTIN.
FT DOMAIN 1214 1272 SUSHI.
FT DISULFID 59 140 BY SIMILARITY.
FT DISULFID 182 253 BY SIMILARITY.
FT DISULFID 206 227 BY SIMILARITY.
FT DISULFID 280 355 BY SIMILARITY.
FT DISULFID 304 325 BY SIMILARITY.
FT DISULFID 1012 1023 BY SIMILARITY.
FT DISULFID 1017 1032 BY SIMILARITY.
FT DISULFID 1034 1043 BY SIMILARITY.
FT DISULFID 1088 1099 BY SIMILARITY.
FT DISULFID 1116 1208 BY SIMILARITY.
FT DISULFID 1184 1200 BY SIMILARITY.
FT DISULFID 1215 1258 BY SIMILARITY.
FT DISULFID 1244 1271 BY SIMILARITY.
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1026 1026 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1223 1223 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 1254 1254 V -> A (in dbSNP:1064389).
FT FT FTid=VAR_016176.
FT CONFLICT 1234 1234 N -> Y (IN REF. 2).
FT CONFLICT 1282 1282 G -> R (IN REF. 2).
SQ SEQUENCE 1321 AA; 142972 MW; 2BF47F823DB980B8 CRC64;

Query Match 3.8%; Score 129.5; DB 1; Length 1321;
Best Local Similarity 22.0%; Pred. No. 2; Indels 201; Gaps 32;
Matches 140; Conservative 63; Mismatches 231

QY 22 PGQSGRRRRGRRSGSGGQF-----WGDRVDSQP-----FAIPYIHPTNPFAPDVT 67
DB 472 PTDPMERRGRFKGLNGRYFQQQPEPEGLQGMESAQPPTSBAVNVQMEP--PLAMAVT 529
QY 68 AAGAGPRVRQPARPLGSAWRD---QAORPAVASRRRPTTAGAA-----PLTAVAPA 116
DB 530 EMLGSG-QSRSP-----WADLTNEYDMFGAGS-----AGGKSSPEPWLWPTTVPVS 575
QY 117 ---HDTPPVPDVP-SRGAILR-RQYNLTSTPLTSSVATGTLNLYAAPLS-----PLLPL 166
DB 576 ISGSRAPULELEKAGPSARPATPDLFWSPLEATVSAPAPWEAFPPVATSDPLPMAM 635
QY 167 QDGTNTHIM-----ATEASNAQYVARATIRYRPLVPNAVGGYIAISIFWFQTTTPT 220
DB 636 LRGPKEWMLPHPTPISTEANRVEAHEATATAPSPAETKV--YSLPLSL-----TPT 687
QY 221 SVMDSITSDVILLVQPGIASLIVPSERLHYRNQWRSVETSGVAEEAETS--GLVML 278
DB 688 G-----QGGEMPTTPESPADFRFETGETSPAQVNKAHSSSSPWPVSNR 732
QY 279 CIHGLSVNTPTPYTGALGLDFALEPRNLTPGNTNTRVSRYSSTARHRLRRGADGT 338
DB 733 NVAVGFPVTETATEPTGLRGI-----PGSES-----GVFDT 763
QY 339 AEUTTTAATFMKDLV-----FTSTNGVGEIGR-GIALT----- 371

Search completed: August 16, 2004, 13:34:13
Job time : 17 secs

Db 764 AESPTSGLQATVDEVDQPMFSPVYVKGLDASSPSAPLGSPGVFLVPKVTNPNLEPWATDEG 823
QY 372 -LFNLADTLLGGLPTL---ISSAGGQLFYSRPVVSANGEPYVKLYTSV-----ENAOQDKG 424
DB 824 PTVPNPMDSVTVPAPSDASGIWEPGQVF--EEAESTTLPQVALDTSIVTPLTTLEQGDK 881
QY 425 IAIPIHDIDIGESRVVIQDYDNQHEQDRPTPSP-----APSRPFVLRANDVL 471
DB 882 VGVPMSTLGS-----SSQHPPEPDQVETQGTSGASVPPHQSSPLGKEAVP 929
QY 472 WLSLTAAYEQDQSTYSGSTGPFVYSD--SVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQ 530
DB 930 PGPTTAASVGEASASVSSGEPVWDPFSSTLLPVTLGIE-----DPELEVLAGSP----- 978
QY 531 QYSKTFVFLPLRKLKLSFWE---AGITKA--GYPN 560
DB 979 -----GVESFWEVASGEEPALPGTPMN 1001

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 13:24:37 ; Search time 42 Seconds
(without alignments)
4958.143 Million cell updates/sec

Title: US-09-851-410A-8

Perfect score: 3408

Sequence: 1 MRPRILLLLMLPMLPAP.....QSTVAELQLKMKVKTREL 660

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertibrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3392	99.5	660	12	Q9E8G5	Q9E8G5 hepatitis e
2	3385	99.3	660	12	Q91855	Q91855 hepatitis e
3	3383	99.3	660	12	Q9WQAO	Q9WQAO hepatitis e
4	3382	99.2	660	12	Q89468	Q89468 hepatitis e
5	3382	99.2	660	12	Q81871	Q81871 hepatitis e
6	3381	99.2	660	12	Q9W7W9	Q9W7W9 hepatitis e
7	3374	99.0	660	12	Q9419	Q9419 hepatitis e
8	3372	98.9	660	12	Q81878	Q81878 hepatitis e
9	3367	98.8	660	12	Q91856	Q91856 hepatitis e
10	3367	98.8	660	12	Q9411	Q9411 hepatitis e
11	3362	98.7	660	12	Q9WL14	Q9WL14 hepatitis e
12	3342.5	98.1	659	12	Q8JT32	Q8JT32 hepatitis e
13	3336	97.9	660	12	Q88985	Q88985 hepatitis e
14	3379	93.3	674	12	Q8JUM6	Q8JUM6 hepatitis e
15	3176	93.2	660	12	Q8JUM8	Q8JUM8 hepatitis e
16	3175	93.2	660	12	Q91114	Q91114 hepatitis e

17	3174	93.1	660	12	Q8JUM4	Q8JUM4 hepatitis e
18	3174	93.1	660	12	Q8AZM3	Q8AZM3 hepatitis e
19	3172	93.1	660	12	Q9YLQ9	Q9YLQ9 hepatitis e
20	3171	93.0	660	12	Q9YLK2	Q9YLK2 hepatitis e
21	3167	92.9	671	12	Q8JUM1	Q8JUM1 hepatitis e
22	3166	92.9	660	12	Q8V729	Q8V729 swine hepat
23	3166	92.9	674	12	Q8JUM2	Q8JUM2 hepatitis e
24	3163	92.8	660	12	Q36613	Q36613 swine hepat
25	3160	92.7	674	12	Q806E0	Q806E0 swine hepat
26	3160	92.7	674	12	Q806D7	Q806D7 hepatitis e
27	3160	92.7	674	12	Q7TGF1	Q7TGF1 hepatitis e
28	3159	92.7	660	12	Q8BB05	Q8BB05 swine hepat
29	3158	92.7	674	12	Q80IR5	Q80IR5 hepatitis e
30	3157	92.6	660	12	Q80IY5	Q80IY5 hepatitis e
31	3153	92.5	660	12	Q80IR7	Q80IR7 hepatitis e
32	3103	91.1	672	12	Q9IVZ8	Q9IVZ8 hepatitis e
33	3053	89.6	605	12	Q9WLO	Q9WLO hepatitis e
34	2688	78.9	550	12	Q8JVV3	Q8JVV3 hepatitis e
35	2660	78.1	525	12	Q39947	Q39947 hepatitis e
36	2381	69.9	486	12	Q9WNN2	Q9WNN2 hepatitis e
37	2298	67.4	466	12	Q9WNN1	Q9WNN1 hepatitis e
38	2197	64.5	436	12	Q9W148	Q9W148 hepatitis e
39	1762	51.7	344	12	Q9WJZ5	Q9WJZ5 hepatitis e
40	1416	41.5	286	12	Q8JQ92	Q8JQ92 swine hepat
41	1411	41.4	286	12	Q8JPR3	Q8JPR3 swine hepat
42	1409	41.3	286	12	Q8JQ93	Q8JQ93 swine hepat
43	1405	41.2	283	12	Q9YEB1	Q9YEB1 hepatitis e
44	1334.5	39.2	606	12	Q913Y7	Q913Y7 avian hepat
45	1320	38.7	258	12	Q81861	Q81861 hepatitis e

ALIGNMENTS

RESULT 1

Q9E8G5 PRELIMINARY; PRT; 660 AA.
 AC Q9E8G5;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Structural protein.
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Abb-2B;
 RX MEDLINE=20271579; PubMed=10813471;
 RA van Cuyck-Gandre H., Zhang H.Y., Tsarev S.A., Warren R.L.,
 RA Caudill J.D., Snellings N.J., Begot L., Innis B.L., Longer C.F.,
 RT "Phylogenetically distinct hepatitis E viruses in Pakistan."
 RL Am. J. Trop. Med. Hyg. 62:187-189(2000).
 DR EMBL; AF185822; AAG16766.1;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR004261; SP2.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF03014; SP2; 1.
 SQ SEQUENCE 660 AA; 70903 MW; 1F506BE3CFB3BACE CRC64;

Query Match 99.5%; Score 3392; DB 12; Length 660;

Best Local Similarity 99.5%; Pred. No. 1.2e-213; Indels 0; Gaps 0;

Matches 657; Conservative 0; Mismatches 3;

Qy 1 MRPRILLLLMLPMLPAPPPGSGRRRRSGSGGFGWGDVDSQPPAIPYIHTN 60

Db 1 MRPRILLLLMLPMLPAPPPGSGRRRRSGSGGFGWGDVDSQPPAIPYIHTN 60

Qy 61 PFAPDVTAAAGAGPRVRQPARPLGSANWDQQRPAVASRRRPTTAGAAPLTAAPAHOTP 120

Db 61 PFAPDVTAAAGAGPRVRQPARPLGSANWDQQRPAVASRRRPTTAGAAPLTAAPAHOTP 120

```
QY 121 PVPDVSRAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAISSIFWPTQTTTPTSDVMNSITSDVRLVQPGI 240
Db 181 NYAQYVARATIRYRPLVNAVGGYAISSIFWPTQTTTPTSDVMNSITSDVRLVQPGI 240
QY 241 ASLVIPSPERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSFVNSYNTPTYTGAIGLL 300
Db 241 ASLVIPSPERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSFVNSYNTPTYTGAIGLL 300
QY 301 DFALIEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALIEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLTLLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLTLLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSPFVLRANDVLMVLSLTAAY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSPFVLRANDVLMVLSLTAAY 480
QY 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDRPLSTIQOYKSTFFVLP 540
Db 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDRPLSTIQOYKSTFFVLP 540
QY 541 LRKGLSFWAGTTKAGYPYNYNTTASDQLLVENAAAGHVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWAGTTKAGYPYNYNTTASDQLLVENAAAGHVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDLTDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660
Db 601 LAPHSALELLEDLTDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660
```

RESULT 2

```
Q91855 PRELIMINARY; PRT; 660 AA.
ID Q91855
AC Q91855;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=93-Egypt;
RX MEDLINE=99105430; PubMed=9890424;
RA Tsarev S.A., Binn L.N., Gomatos P.J., Arthur R.R., Monier M.K.,
RA van Cuyck-Gandre H., Longer C.F., Innis B.L.;
RT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
RL J. Med. Virol. 57:68-74(1999).
DR ENBL; AF051351; AAC35761.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71040 MW; 044FF5EA7C492791 CRC64;
```

Query Match 99.3%; Score 3385; DB 12; Length 660;
Best Local Similarity 99.1%; Pred. No. 3.4e-213;
Matches 654; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 MRPRPILLLMLFPLMLPAPPQPSGRRRSGSGGFGWGDVDSQPFALPYIHTPTN 60
Db 1 MRPRPILLLMLFPLMLPAPPQPSGRRRSGSGGFGWGDVDSQPFALPYIHTPTN 60
QY 61 PFAPDVTAAAGPRVRQPARPLGSAWRDQAQRPAVARRRPTTAGAAPLTAVAPADHTP 120
```

```
Db 61 PFAPDVTAAAGPRVRQPARPLGSAWRDQAQRPAVARRRPTTAGAAPLTAVAPADHTP 120
QY 121 PVPDVSRAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRAILRRQYNLSPLTSSVATGNTLVLYAAPLSPLPLQDGTNTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAISSIFWPTQTTTPTSDVMNSITSDVRLVQPGI 240
Db 181 NYAQYVARATIRYRPLVNAVGGYAISSIFWPTQTTTPTSDVMNSITSDVRLVQPGI 240
QY 241 ASLVIPSPERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSFVNSYNTPTYTGAIGLL 300
Db 241 ASLVIPSPERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSFVNSYNTPTYTGAIGLL 300
QY 301 DFALIEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALIEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLTLLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLTLLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSPFVLRANDVLMVLSLTAAY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQEDRPTSPAPSPFVLRANDVLMVLSLTAAY 480
QY 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDRPLSTIQOYKSTFFVLP 540
Db 481 DQSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDRPLSTIQOYKSTFFVLP 540
QY 541 LRKGLSFWAGTTKAGYPYNYNTTASDQLLVENAAAGHVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWAGTTKAGYPYNYNTTASDQLLVENAAAGHVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDLTDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660
Db 601 LAPHSALELLEDLTDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTREL 660
```

RESULT 3

```
Q9WQAO PRELIMINARY; PRT; 660 AA.
ID Q9WQAO
AC Q9WQAO;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF-2.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AKL-90;
RX MEDLINE=99350000; PubMed=10423137;
RA Arankalle V.A., Paranjape S., Emerson S.U., Purcell R.H.,
RA Walimbe A.M.;
RT "Phylogenetic analysis of hepatitis E virus isolates from India (1976-1993).";
RL J. Gen. Virol. 80:1691-1700(1999).
DR EMBL; AF124407; AAD45493.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71041 MW; 03B72DDF0AB7B521 CRC64;
```

Query Match 99.3%; Score 3383; DB 12; Length 660;
Best Local Similarity 99.2%; Pred. No. 4.6e-213;
Matches 655; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 MRPRPILLLMLFPLMLPAPPQPSGRRRSGSGGFGWGDVDSQPFALPYIHTPTN 60
```

```
Db 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60
Qy 61 PFAPDVTAAAGPVRQPARPLGSAWRDQAPVASSRRRTTGAAPLTAAPAHDTP 120
Db 61 PFAPDVTAAAGPVRQPARPLGSAWRDQAPVASSRRRTTGAAPLTAAPAHDTP 120
Qy 121 PVPDVSAGAILRRQYNLSTSLTSSVATGTLNLYAAPLSPLLLPLOGDGTNTHIMATEAS 180
Db 121 PVPDVSAGAILRRQYNLSTSLTSSVATGTLNLYAAPLSPLLLPLOGDGTNTHIMATEAS 180
Qy 181 NYAQYRVARATIRYRPLVPNAVGGYAISISFWPQTTTPTSDVMNSITSTDVRIILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVPNAVGGYAISISFWPQTTTPTSDVMNSITSTDVRIILVQPGI 240
Qy 241 ASELVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSPVNSYNTPTYTGA LGLL 300
Db 241 ASELVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSPVNSYNTPTYTGA LGLL 300
Qy 301 DFALSEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTTAAATRFMKDLYFTSTNG 360
Db 301 DFALSEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTTAAATRFMKDLYFTSTNG 360
Qy 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Qy 421 QDKGIAIPHIDILGESRVVIQDYNQHEQDRTPSPAPSRPFSVLRANDVWLSLTAABY 480
Db 421 QDKGIAIPHIDILGESRVVIQDYNQHEQDRTPSPAPSRPFSVLRANDVWLSLTAABY 480
Qy 481 DOSTYSSGTPVYVSDSVTLNVNATGAQAVASRLDWTKTLDGRPLSTTQQYKTFVFLP 540
Db 481 DOSTYSSGTPVYVSDSVTLNVNATGAQAVASRLDWTKTLDGRPLSTTQQYKTFVFLP 540
Qy 541 LRKLSFWEAGTTKAGYPYNNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWEAGTTKAGYPYNNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Qy 601 LAPHASALLETLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKTRREL 660
Db 601 LAPHASALLETLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKTRREL 660

RESULT 4
Q89468 PRELIMINARY; PRT; 660 AA.
AC Q89468;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE UNNAMED protein product.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]_
RP SEQUENCE FROM N.A.
RA Yin S.R., Purcell R.H., Emerson S.U.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=K52-87;
RX MEDLINE=95176571; PubMed=7871758;
RA Yin S., Purcell R.H., Emerson S.U.;
RT "A new Chinese isolate of hepatitis E virus: comparison with strains
RT recovered from different geographical regions.";
RL Virus Genes 9:23-32(1994).
DR EMBL; L25547; AAA31080.1; -.
DR EMBL; L25595; AAA65490.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
```

```
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70979 MW; DC68116DCD639175 CRC64;

Query Match 99.2%; Score 3382; DB 12; Length 660;
Best Local Similarity 99.1%; Pred. No. 5.3e-213;
Matches 654; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60
Db 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60
Qy 61 PFAPDVTAAAGPVRQPARPLGSAWRDQAPVASSRRRTTGAAPLTAAPAHDTP 120
Db 61 PFAPDVTAAAGPVRQPARPLGSAWRDQAPVASSRRRTTGAAPLTAAPAHDTP 120
Qy 121 PVPDVSAGAILRRQYNLSTSLTSSVATGTLNLYAAPLSPLLLPLOGDGTNTHIMATEAS 180
Db 121 PVPDVSAGAILRRQYNLSTSLTSSVATGTLNLYAAPLSPLLLPLOGDGTNTHIMATEAS 180
Qy 181 NYAQYRVARATIRYRPLVPNAVGGYAISISFWPQTTTPTSDVMNSITSTDVRIILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVPNAVGGYAISISFWPQTTTPTSDVMNSITSTDVRIILVQPGI 240
Qy 241 ASELVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSPVNSYNTPTYTGA LGLL 300
Db 241 ASELVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSPVNSYNTPTYTGA LGLL 300
Qy 301 DFALSEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTTAAATRFMKDLYFTSTNG 360
Db 301 DFALSEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTTAAATRFMKDLYFTSTNG 360
Qy 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Qy 421 QDKGIAIPHIDILGESRVVIQDYNQHEQDRTPSPAPSRPFSVLRANDVWLSLTAABY 480
Db 421 QDKGIAIPHIDILGESRVVIQDYNQHEQDRTPSPAPSRPFSVLRANDVWLSLTAABY 480
Qy 481 DOSTYSSGTPVYVSDSVTLNVNATGAQAVASRLDWTKTLDGRPLSTTQQYKTFVFLP 540
Db 481 DOSTYSSGTPVYVSDSVTLNVNATGAQAVASRLDWTKTLDGRPLSTTQQYKTFVFLP 540
Qy 541 LRKLSFWEAGTTKAGYPYNNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWEAGTTKAGYPYNNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Qy 601 LAPHASALLETLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKTRREL 660
Db 601 LAPHASALLETLDYPARAHTFDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKTRREL 660

RESULT 5
Q81871 PRELIMINARY; PRT; 660 AA.
AC Q81871;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE ORF 2 precursor.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
RN [2]
```

RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Rikihisa M., Rikihisa T., Mizuno K., Win K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92335008; PubMed=1630924;
RA Aye T.T., Uchida T., Ma X.Z., Iida F., Shikata T., Zhuang H.,
RA Win K.M.;
RT "Complete nucleotide sequence of a hepatitis E virus isolated from the
RT Xinjiang epidemic (1986-1988) of China.";
RL Nucleic Acids Res. 20:3512-3512(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.U., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271462; PubMed=1589964;
RA Fry K.E., Tam A.W., Smith M.W., Kim J.P., Luk K.-C., Young L.M.,
RA Piatk M., Feldman R.A., Yun K.Y., Purdy M.A., McCaustland K.A.,
RA Bradley D.W., Reyes G.R.;
RT "Hepatitis E virus (HEV): Strain variation in the nonstructural gene
RT region encoding consensus motifs for an RNA-dependent RNA polymerase
RT and an ATP/GTP binding site.";
RL Virus Genes 6:173-185(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93348763; PubMed=9346669;
RA Bi S.-L., Purdy M.A., McCaustland K.A., Margolis H.S., Bradley D.W.;
RT "The sequence of hepatitis E virus isolated directly from a single
RT source during an outbreak in China.";
RL Virus Res. 28:233-247(1993).
DR EMBL: L08816; AAA03191.1; -;
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
KW Signal.
FT SIGNAL.
FT SIGNAL.
SQ SEQUENCE 1 19 POTENTIAL.
Query Match 1 19 660 AA; 70979 MW; B58F23955FDD6614 CRC64;
Best Local Similarity 99.2%; Score 3382; DB 12; Length 660;
Matches 654; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRGRSGSGGFGWDRVDSQPFALPYIHTN 60
DB 1 MRPRPILLLLMLPMLPAPPQPSGRRRGRSGSGGFGWDRVDSQPFALPYIHTN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSARWDOAQPAAVARRRPTTAGAAPTAVAPADTP 120
DB 61 PFAPDVTAAGAGPRVQPARPLGSARWDOAQPAAVARRRPTTAGAAPTAVAPADTP 120
QY 121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLLPLODGTNTHIMATEAS 180
DB 121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGYAISISFWQTTTTPTSVDMNSTTSDVRLVQPGI 240
DB 181 NYAQYRVARATIRYRPLVNAVGYAISISFWQTTTTPTSVDMNSTTSDVRLVQPGI 240
QY 241 ASELVIPSERLHYNQWRSVETSGVAEEATSGLVMLCHIGSLVNSYNTPTGTALGLL 300
DB 241 ASELVIPSERLHYNQWRSVETSGVAEEATSGLVMLCHIGSLVNSYNTPTGTALGLL 300
QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360

DB 301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLNADTLGLGLPTELISSAGGOLFYSRPPVSVANGSPPTVKLYTSVNAQ 420
DB 361 VGEIGRGIALTLNADTLGLGLPTELISSAGGOLFYSRPPVSVANGSPPTVKLYTSVNAQ 420
QY 421 QDKGIAIPHDIDIGESRVVIQDYDNOHQDRPTSPAPSRPFVLRANDVLMSLTAAEY 480
DB 421 QDKGIAIPHDIDIGESRVVIQDYDNOHQDRPTSPAPSRPFVLRANDVLMSLTAAEY 480
QY 481 DQSTYSSSTGPVVSVDSTLVNVTGAQAVARSIDWTKVLDGRPLSTIQOYSKTFVLP 540
DB 481 DQSTYSSSTGPVVSVDSTLVNVTGAQAVARSIDWTKVLDGRPLSTIQOYSKTFVLP 540
QY 541 LRKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTVTTSLGAGPVSISAV 600
DB 541 LRKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTVTTSLGAGPVSISAV 600
QY 601 LAPHSALLLEDITLDYPARAHTFDDFCPECRPLGLOGCAFQSTVAELQRLKMKVKTREL 660
DB 601 LAPHSALLLEDITLDYPARAHTFDDFCPECRPLGLOGCAFQSTVAELQRLKMKVKTREL 660
RESULT 6
Q9W7W9 PRELLIMINARY; PRT; 660 AA.
AC Q9W7W9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Structural protein 2 (Structural viral protein).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HEV-Morocco;
RX MEDLINE=99107414; PubMed=9892396;
RA Meng J., Cong M., Dai X., Pillot J., Purdy M.A., Fields H.A.,
RA Khudyakov Y.E.;
RT "Primary structure of open reading frame 2 and 3 of the hepatitis E
RT virus isolated from Morocco.";
RL J. Med. Virol. 57:126-133(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Morocco;
RA Chen G., Meng J.;
RT "Identification of the 5' Capped and 3' Complete Terminal Sequence of
RT the Hepatitis E virus Isolated from Morocco.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF065061; AAD10627.1; -;
DR EMBL: AY230302; AAO72992.1; -;
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70943 MW; 9719D30CCD932950 CRC64;
Query Match 99.2%; Score 3381; DB 12; Length 660;
Best Local Similarity 99.2%; Pred. No. 6.2e-213;
Matches 655; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRGRSGSGGFGWDRVDSQPFALPYIHTN 60
DB 1 MRPRPILLLLMLPMLPAPPQPSGRRRGRSGSGGFGWDRVDSQPFALPYIHTN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSARWDOAQPAAVARRRPTTAGAAPTAVAPADTP 120
DB 61 PFAPDVTAAGAGPRVQPARPLGSARWDOAQPAAVARRRPTTAGAAPTAVAPADTP 120
QY 121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLLPLODGTNTHIMATEAS 180
DB 121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLLPLODGTNTHIMATEAS 180


```

Db 121 PVPDVSRAILRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPLQDGTNTHIMATEAS 180
QY 181 NVAQYVARATIRYRLPVNAVGGYAISISFWPQTITPTSDVMSITSDVRLVQPGI 240
Db 181 NVAQYVARATIRYRLPVNAVGGYAISISFWPQTITPTSDVMSITSDVRLVQPGI 240
QY 241 ASSELVPSRLHYRNQWRSVETSGVAEBEATSGLVMLCIHGSVNSYNTPTGALGLL 300
Db 241 ASSELVPSRLHYRNQWRSVETSGVAEBEATSGLVMLCIHGSVNSYNTPTGALGLL 300
QY 301 DFALELEFNLPNGTNTVRYSYSTARHLRKGADGTAELTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFNLPNGTNTVRYSYSTARHLRKGADGTAELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNQHQDRPTSPAPSRPFSVLRANDVLMSLTAAY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNQHQDRPTSPAPSRPFSVLRANDVLMSLTAAY 480
QY 481 DOSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKTLDGRPLSTIQOYSKTFVLP 540
Db 481 DOSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKTLDGRPLSTIQOYSKTFVLP 540
QY 541 LRKLSFWAGTTKAGYPYNYNTASDQLLVENAGHRVAISYTTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWAGTTKAGYPYNYNTASDQLLVENAGHRVAISYTTTSLGAGPVSISAVAV 600
QY 601 LAPHSALEDTLDYPARAHTFDDFCPCPRPLGQCAFQSTVABLQRLKMKVGTREL 660
Db 601 LAPHSALEDTLDYPARAHTFDDFCPCPRPLGQCAFQSTVABLQRLKMKVGTREL 660

RESULT 7
Q69419
ID Q69419 PRELIMINARY; PRT; 660 AA.
AC Q69419;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE ORF1, ORF2 & ORF3.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RA von Brunn A., Seebach J., Thyagarajan S.P., Mohanavalli B., Menon T.,
RA Froesner G.;
RT "PCR amplification, cloning and sequence determination of a hepatitis
RT E virus isolate from Madras, India.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X99441; CAA67804.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70960 MW; DAC3DF95F91689F8 CRC64;

```

```

Query Match      99.0%; Score 3374; DB 12; Length 660;
Best Local Similarity 99.1%; Pred. No. 1.8e-212;
Matches 654; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRRSGSGGFGWDRVDSQFFAIPYIHPTN 60
Db 1 MRPRPILLLLMLPMLPAPPQPSGRRRRSGSGGFGWDRVDSQFFAIPYIHPTN 60
QY 61 PPAPDVTAAGAGPRVQPARPLGSAWRDQARPAVASRRRPTTAGAAPTAVPAHDTF 120
Db 61 PPAPDVTAAGAGPRVQPARPLGSAWRDQARPAVASRRRPTTAGAAPTAVPAHDTF 120

```

```

QY 121 PVPDVSRAILRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPLQDGTNTHIMATEAS 180
Db 121 PVPDVSRAILRQYNLSTPLTSSVATGNTLVLYAAPLSPLLPLQDGTNTHIMATEAS 180
QY 181 NVAQYVARATIRYRLPVNAVGGYAISISFWPQTITPTSDVMSITSDVRLVQPGI 240
Db 181 NVAQYVARATIRYRLPVNAVGGYAISISFWPQTITPTSDVMSITSDVRLVQPGI 240
QY 241 ASSELVPSRLHYRNQWRSVETSGVAEBEATSGLVMLCIHGSVNSYNTPTGALGLL 300
Db 241 ASSELVPSRLHYRNQWRSVETSGVAEBEATSGLVMLCIHGSVNSYNTPTGALGLL 300
QY 301 DFALELEFNLPNGTNTVRYSYSTARHLRKGADGTAELTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFNLPNGTNTVRYSYSTARHLRKGADGTAELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNQHQDRPTSPAPSRPFSVLRANDVLMSLTAAY 480
Db 421 QDKGIAIPHDIDLGESRVVIQDYDNQHQDRPTSPAPSRPFSVLRANDVLMSLTAAY 480
QY 481 DOSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKTLDGRPLSTIQOYSKTFVLP 540
Db 481 DOSTYSSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKTLDGRPLSTIQOYSKTFVLP 540
QY 541 LRKLSFWAGTTKAGYPYNYNTASDQLLVENAGHRVAISYTTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWAGTTKAGYPYNYNTASDQLLVENAGHRVAISYTTTSLGAGPVSISAVAV 600
QY 601 LAPHSALEDTLDYPARAHTFDDFCPCPRPLGQCAFQSTVABLQRLKMKVGTREL 660
Db 601 LAPHSALEDTLDYPARAHTFDDFCPCPRPLGQCAFQSTVABLQRLKMKVGTREL 660

RESULT 8
Q81878
ID Q81878 PRELIMINARY; PRT; 660 AA.
AC Q81878;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Complete genome sequence.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Uighi79;
RA Uchida T.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; D11093; BAA01867.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70937 MW; 5412313F1A2EF4A9 CRC64;

```

```

Query Match      98.9%; Score 3372; DB 12; Length 660;
Best Local Similarity 98.6%; Pred. No. 2.4e-212;
Matches 651; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRRSGSGGFGWDRVDSQFFAIPYIHPTN 60
Db 1 MRPRPILLLLMLPMLPAPPQPSGRRRRSGSGGFGWDRVDSQFFAIPYIHPTN 60
QY 61 PPAPDVTAAGAGPRVQPARPLGSAWRDQARPAVASRRRPTTAGAAPTAVPAHDTF 120
Db 61 PPAPDVTAAGAGPRVQPARPLGSAWRDQARPAVASRRRPTTAGAAPTAVPAHDTF 120

```

```

QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVAAPLSPLLPQDGTNTTHIMATEAS 180
DB 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVAAPLSPLLPQDGTNTTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTTTTPTSDVMSITSDVRLVQPGI 240
DB 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTTTTPTSDVMSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSFVNSYNTPTGTALGLL 300
DB 241 ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSFVNSYNTPTGTALGLL 300
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRWVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
DB 421 QDKGIAIPHDIDLGESRWVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIOQYKTFVLP 540
DB 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIOQYKTFVLP 540
QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAAHRAVIAISTVTTSLGAGPVSISAVAV 600
DB 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAAHRAVIAISTVTTSLGAGPVSISAVAV 600
QY 601 LAPHSAALLEDLTDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660
DB 601 LAPHSAALLEDLTDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660

RESULT 9
O91856 PRELIMINARY; PRT; 660 AA.
AC O91856;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=94-Egypt;
RX MEDLINE=99105430; PubMed=9890424;
RA Tsarev S.A., Binn L.N., Gomatos P.J., Arthur R.R., Monier M.K.,
RA van Cuyck-Gandre H., Longer C.F., Innis B.L.;
RT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
RL J. Med. Virol. 57:68-74(1999).
DR EMBL; AF051352; AAC35764.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71026 MW; 16C560FA16941F2A CRC64;

```

```

Query Match 98.8%; Score 3367; DB 12; Length 660;
Best Local Similarity 98.8%; Pred. No. 5.1e-212;
Matches 652; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 MRPRPILLLLMLPMLPAPPGPSPGRRRGGSGGPGWGRVDSQPFAIPYIHTN 60
DB 1 MRPRPILLLLMLPMLPAPPGPSPGRRRGGSGGPGWGRVDSQPFAIPYIHTN 60
QY 61 PFAPDVTAAAGAGPRVRQPRPLGSARWDAQRPVAVSRRPRTTAGAAPTAVAPAHDT 120

```

```

DB 61 PFAPDVTAAAGAGPRVRQPRPLGSARWDAQRPVAVSRRPRTTAGAAPTAVAPAHDT 120
QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVAAPLSPLLPQDGTNTTHIMATEAS 180
DB 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNVLVAAPLSPLLPQDGTNTTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTTTTPTSDVMSITSDVRLVQPGI 240
DB 181 NYAQYVARATIRYRPLVNAVGGYAIISFWPQTTTTPTSDVMSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSFVNSYNTPTGTALGLL 300
DB 241 ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSFVNSYNTPTGTALGLL 300
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRWVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
DB 421 QDKGIAIPHDIDLGESRWVIQDYDNOHQEDRPTSPAPSRPFSVLRANDVWLSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIOQYKTFVLP 540
DB 481 DQSTYGSSTGPVYVSDSVTLVNVATGAQAVARSLDWTKVLDGRPLSTIOQYKTFVLP 540
QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAAHRAVIAISTVTTSLGAGPVSISAVAV 600
DB 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAAHRAVIAISTVTTSLGAGPVSISAVAV 600
QY 601 LAPHSAALLEDLTDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660
DB 601 LAPHSAALLEDLTDYPARAHTFDDFCPCRPLGLOGCAFQSTVAELQRLKMKVGTREL 660

RESULT 10
Q69411 PRELIMINARY; PRT; 660 AA.
AC Q69411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=hev037;
RX Donati M.C., Fagan E.A., Harrison T.J.;
RT "Sequence analysis of full length HEV clones derived directly from
RT human liver in fulminant hepatitis E.";
RL (In) Rizzetto M., Purcell R.H., Gerin J.L., Verme G (eds.);
RL VIRAL HEPATITIS AND LIVER DISEASE, pp.313-316,
RL Edizioni Minerva Medica, Torino (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=hev037;
RX Harrison T.J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X98292; CAA66937.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71024 MW; F37E3FEFF7A4EAD1 CRC64;

```

```
Query Match      98.8%; Score 3367; DB 12; Length 660;
Best Local Similarity 98.5%; Pred. No. 5.1e-212;
Matches 650; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRPRILLLLLMFLPMLPAPPQSGRRRRSGSGGFGWDRVDSQPPFAIPYIHPN 60
DB 1 MRPRILLLLLMFLPMLPAPPQSGRRRRSGSGGFGWDRVDSQPPFAIPYIHPN 60

QY 61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQORPAVASRRRTTAGAAPLTAVAPADTP 120
DB 61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQORPAASRRRTTAGAAPLTAVAPADTP 120

QY 121 PVPDVDSGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
DB 121 PVPDVDSGAILRRQYNLSTSLTSPVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180

QY 181 NYAQYRVARATIRYRPLVNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIILVQPGI 240
DB 181 NYAQYRVARATIRYRPLVNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIILVQPGI 240

QY 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 300
DB 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 300

QY 421 QDKGIAIHDIIDLGSRVVIQDYDNOHQEODRTPSPAPRPSFVLRANDVLSLTAABY 480
DB 421 QDKGIAIHDIIDLGSRVVIQDYDNOHQEODRTPSPAPRPSFVLRANDVLSLTAABY 480

QY 481 LRGKLSFWEAGTTKAGYPYNYNTTASDQLLENAAGHRVAISTYTTSLGAGPVSISAVAV 600
DB 481 LRGKLSFWEAGTTKAGYPYNYNTTASDQLLENAAGHRVAISTYTTSLGAGPVSISAVAV 600

QY 601 LAPHASALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTR 660
DB 601 LAPHASALLEDTMDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTR 660

RESULT 11
Q9WLL4 PRELIMINARY; PRT; 660 AA.
AC Q9WLL4;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=TK15/92;
RX MEDLINE=99049628; PubMed=9833882;
RA Gouvea V., Snellings N., Popek M.J., Longer C.F., Innis B.L.;
RT "Hepatitis E virus: complete genome sequence and phylogenetic analysis
of a Nepali isolate.";
RL Virus Res. 57:21-26(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TK15/92;
RX MEDLINE=22072932; PubMed=12076829;
RA Jameel S., Zafrullah M., Chawla Y.K., Dillawari J.B.;
```

```
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF051830; AAC97188.1;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70798 MW; A04C0185ACC085DB CRC64;
```

```
Query Match      98.7%; Score 3362; DB 12; Length 660;
Best Local Similarity 98.8%; Pred. No. 1.1e-211;
Matches 652; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 MRPRILLLLLMFLPMLPAPPQSGRRRRSGSGGFGWDRVDSQPPFAIPYIHPN 60
DB 1 MRPRILLLLLMFLPMLPAPPQSGRRRRSGSGGFGWDRVDSQPPFAIPYIHPN 60

QY 61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQORPAVASRRRTTAGAAPLTAVAPADTP 120
DB 61 PPAPDVTAAGAGPRVROPAPPLGSAWRDQORPAASRRRTTAGAAPLTAVAPADTP 120

QY 121 PVPDVDSGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
DB 121 PVPDVDSGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180

QY 181 NYAQYRVARATIRYRPLVNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIILVQPGI 240
DB 181 NYAQYRVARATIRYRPLVNAVGGVAISISFWPQTITPTSDVMNSITSTDVRIILVQPGI 240

QY 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 300
DB 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTYGALGLL 300

QY 301 DFALSEFRNLTPGNTNTRVSYSTARHRLRGADGTAEITTTAATREMKDLYFTSTNG 360
DB 301 DFALSEFRNLTPGNTNTRVSYSTARHRLRGADGTAEITTTAATREMKDLYFTSTNG 360

QY 361 VGEIORGIALTLFNLADTLGLLPTLTISSAGGQLFYSPRPVSANGEPVTKLYTSVENAQ 420
DB 361 VGEIORGIALTLFNLADTLGLLPTLTISSAGGQLFYSPRPVSANGEPVTKLYTSVENAQ 420

QY 421 QDKGIAIHDIIDLGSRVVIQDYDNOHQEODRTPSPAPRPSFVLRANDVLSLTAABY 480
DB 421 QDKGIAIHDIIDLGSRVVIQDYDNOHQEODRTPSPAPRPSFVLRANDVLSLTAABY 480

QY 481 LRGKLSFWEAGTTKAGYPYNYNTTASDQLLENAAGHRVAISTYTTSLGAGPVSISAVAV 600
DB 481 LRGKLSFWEAGTTKAGYPYNYNTTASDQLLENAAGHRVAISTYTTSLGAGPVSISAVAV 600

QY 601 LAPHASALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTR 660
DB 601 LAPHASALLEDTLDYPARAHTFDDFCPCRPGLGQCAFQSTVAELQRLKMKVGKTR 660

RESULT 12
Q8JUT32 PRELIMINARY; PRT; 659 AA.
AC Q8JUT32;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072932; PubMed=12076829;
RA Jameel S., Zafrullah M., Chawla Y.K., Dillawari J.B.;
```

RT "Reevaluation of a North India isolate of hepatitis E virus based on
RT the full-length genomic sequence obtained following long RT-PCR";
RL Virus Res. 86:53-58(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Jameel S., Zafrullah M., Chawla Y.K., Dilawari J.B.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF459438; AAM66330.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 659 AA; 70842 MW; ABCC177BBE6012C8 CRC64;

Query Match 98.18; Score 3342.5; DB 12; Length 659;
Best Local Similarity 98.3%; Pred. No. 2e-210;
Matches 649; Conservative 1; Mismatches 9; Indels 1; Gaps 1;
QY 1 MRPRPILLLLMFLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPIHPTN 60
Db 1 MRPRPILLLLMFLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPIHPTN 59
QY 61 PFAPDVTAAGAGPRVQRPVPLGSAWRDQAQRPAAASRRRPTTAGAAPTAVAPADHTP 120
Db 60 PFAPDVTAAGAGPRVQRPVPLGSAWRDQAQRPAAASRRRPTTAGAAPTAVAPADHTP 119
QY 121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 120 PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 179
QY 181 NYAQYRVARATIRYRPLVNAVGGYAIISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 240
Db 180 NYAQYRVARATIRYRPLVNAVGGYAIISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 239
QY 241 ASELVTPSERLHYRNQWRVETSGVAEEATSGLVMLCIHGSIVNSYTNTPYTGALL 300
Db 240 ASELVTPSERLHYRNQWRVETSGVAEEATSGLVMLCIHGSIVNSYTNTPYTGALL 299
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 300 DFALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 359
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPYTKLYTSVENAQ 420
Db 360 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPYTKLYTSVENAQ 419
QY 421 QDKGIALPHDIDGESRVITQDYNQHEQDRPTSPAPSPFVSRLANDVWLSLTAAY 480
Db 420 QDKGIALPHDIDGESRVITQDYNQHEQDRPTSPAPSPFVSRLANDVWLSLTAAY 479
QY 481 DQSTYGSSTGPVVVSDSVTLNVNATGAQAVARSLDWTKVLDGRPLSTIQOYSKTFVLP 540
Db 480 DQSTYGSSTGPVVVSDSVTLNVNATGAQAVARSLDWTKVLDGRPLSTIQOYSKTFVLP 539
QY 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVV 600
Db 540 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVV 599
QY 601 LAPHSALELLEDTLDPARAHFTDDFCPCRPLGLQCAFQSTVAELQRLKMKVKTKREL 660
Db 600 LAPHSALELLEDTLDPARAHFTDDFCPCRPLGLQCAFQSTVAELQRLKMKVKTKREL 659

RESULT 13

Q68985 ID Q68985 PRELIMINARY; PRT; 660 AA.
AC Q68985;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF-2.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;

OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HYDERABAD;
RX MEDLINE=96087204; PubMed=8567900;
RA Panda S.K., Nanda S.K., Zafrullah M., Ansari I.H., Ozdener M.H.,
RA Jameel S.;
RT "An Indian strain of hepatitis E virus (HEV): cloning, sequence, and
RT expression of structural region and antibody responses in sera from
RT individuals from an area of high-level HEV endemicity";
RN J. Clin. Microbiol. 33:2653-2659(1995).
RL [2]
RP SEQUENCE FROM N.A.
RA Ansari I.H., Nanda S.K., Durgapal H., Jameel S., Panda S.K.;
RT "Translation analysis of complete HEV genome";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U22532; AAA97366.1; -
DR EMBL; AF076239; AAC27936.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70936 MW; 948556F804382EB3 CRC64;

Query Match 97.9%; Score 3336; DB 12; Length 660;
Best Local Similarity 98.0%; Pred. No. 5.5e-210;
Matches 647; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMFLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPIHPTN 60
Db 1 MGPRPILLLLMFLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPIHPTN 60
QY 61 PFAPDVTAAGAGPRVQRPVPLGSAWRDQAQRPAAASRRRPTTAGAAPTAVAPADHTP 120
Db 61 PFAPDVTAAGAGPRVQRPVPLGSAWRDQAQRPAAASRRRPTTAGAAPTAVAPADHTP 120
QY 121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSRGAILRRQYNLSPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGGYAIISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGGYAIISFWPQTPTTPTSDVMNSITSTDVRLVQPGI 240
QY 241 ASELVTPSERLHYRNQWRVETSGVAEEATSGLVMLCIHGSIVNSYTNTPYTGALL 300
Db 241 ASELVTPSERLHYRNQWRVETSGVAEEATSGLVMLCIHGSIVNSYTNTPYTGALL 300
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPYTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTELISSAGGOLFYSRPVVSANGEPYTKLYTSVENAQ 420
QY 421 QDKGIALPHDIDGESRVITQDYNQHEQDRPTSPAPSPFVSRLANDVWLSLTAAY 480
Db 421 QDKGIALPHDIDGESRVITQDYNQHEQDRPTSPAPSPFVSRLANDVWLSLTAAY 480
QY 481 DQSTYGSSTGPVVVSDSVTLNVNATGAQAVARSLDWTKVLDGRPLSTIQOYSKTFVLP 540
Db 481 DQSTYGSSTGPVVVSDSVTLNVNATGAQAVARSLDWTKVLDGRPLSTIQOYSKTFVLP 540
QY 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVV 600
Db 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVV 600
QY 601 LAPHSALELLEDTLDPARAHFTDDFCPCRPLGLQCAFQSTVAELQRLKMKVKTKREL 660
Db 601 LAPHSALELLEDTLDPARAHFTDDFCPCRPLGLQCAFQSTVAELQRLKMKVKTKREL 660

```
RESULT 14
Q8JUN6
ID Q8JUN6 PRELIMINARY; PRT; 674 AA.
AC Q8JUN6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ORF2 protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAK-Sai;
RA Mishiro S.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JAK-Sai;
RX MEDLINE=21996248; PubMed=12001054;
RA Takahashi K., Kang J.-H., Ohnishi S., Hino K., Mishiro S.;
RT "Genetic Heterogeneity of Hepatitis E Virus Recovered from Japanese Patients with Acute Sporadic Hepatitis.";
RL J. Infect. Dis. 185:1342-1345(2002).
DR EMBL; AB074915; BAB96557.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 674 AA; 72591 MW; 5C196112F81A2265 CRC64;

Query Match 93.3%; Score 3179; DB 12; Length 674;
Best Local Similarity 92.1%; Pred. No. 1.1e-199;
Matches 607; Conservative 25; Mismatches 27; Indels 0; Gaps 0;

QY 1 MRPRLTLLMLPMLPAPPQPSGRRGRSGSGGFWGDRVDSQFPALPYIHPN 60
DB 15 MRSALLFLFLVFLPMLPAPPQPSGRRGRSGSGGFWGDRVDSQFPALPYIHPN 74

QY 61 PPAVDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADHTP 120
DB 75 PPAVDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADHTP 134

QY 121 PVPDVSRAILLRQYNLTSPVATGNTLVLYAAPLSPLLPQDGTNTTHIMATEAS 180
DB 135 PVPDVSRAILLRQYNLTSPVATGNTLVLYAAPLSPLLPQDGTNTTHIMATEAS 194

QY 181 NYAQVRVATIRYRLVFNAGVYASISFWPQTTTTSVDMNSITSDVRLVQPGI 240
DB 195 NYAQVRVATIRYRLVFNAGVYASISFWPQTTTTSVDMNSITSDVRLVQPGI 254

QY 241 ASELVTPSELHYRNQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGALGLL 300
DB 255 ASELVTPSELHYRNQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGALGLL 314

QY 301 DFALELEFNLTPGNTNTRVSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNG 360
DB 315 DFALELEFNLTPGNTNTRVSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNG 374

QY 361 VGEIGRGIALTLFNLADTLGGPLTELISAGQQLFYSPVVSANGEPTVKLYTSVENAQ 420
DB 375 VGEVGRGIALTLFNLADTLGGPLTELISAGQQLFYSPVVSANGEPTVKLYTSVENAQ 434

QY 421 QDKGIAIPHIDLGESRVVIQDYDQHEQDRPTSPAPSRPFVSLRANDVWLSLTAAY 480
DB 435 QDKGIAIPHIDLGESRVVIQDYDQHEQDRPTSPAPSRPFVSLRANDVWLSLTAAY 494

QY 481 DQSTYSSSTGPPVVSVDVTLVNVATGAQAVARSIDWTKVTLDCRPLSTQQYSKTFVLP 540
DB 495 DQSTYSSSTGPPVVSVDVTLVNVATGAQAVARSIDWTKVTLDCRPLSTQQYSKTFVLP 554

QY 541 LRGLKSFWEAGTTKAGYPYNYNTTASDQLIVENAGHRVAISTYTTTSLGAGPVVISAVAV 600

Db 555 LRGLKSFWEAGTTKAGYPYNYNTTASDQLIVENAGHRVAISTYTTTSLGAGPVVISAVGV 614
QY 601 LAPHSALALLEDTLDYPARAHTFDDFCPECRPLGLOGCAFOSTVAELQRLKMKVGKTR 659
DB 615 LAPHSALALLEDTLDYPARAHTFDDFCPECRPLGLOGCAFOSTVAELQRLKMKVGKTR 673

RESULT 15
Q8JUM8
ID Q8JUM8 PRELIMINARY; PRT; 660 AA.
AC Q8JUM8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF2 protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JKN-Sap;
RA Mishiro S.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JKN-Sap;
RX MEDLINE=21996248; PubMed=12001054;
RA Takahashi K., Kang J.-H., Ohnishi S., Hino K., Mishiro S.;
RT "Genetic Heterogeneity of Hepatitis E Virus Recovered from Japanese Patients with Acute Sporadic Hepatitis.";
RL J. Infect. Dis. 185:1342-1345(2002).
DR EMBL; AB074918; BAB96562.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70980 MW; 19233C22F3E57FF8 CRC64;

Query Match 93.2%; Score 3176; DB 12; Length 660;
Best Local Similarity 92.1%; Pred. No. 1.6e-199;
Matches 607; Conservative 26; Mismatches 26; Indels 0; Gaps 0;

QY 1 MRPRLTLLMLPMLPAPPQPSGRRGRSGSGGFWGDRVDSQFPALPYIHPN 60
DB 1 MRPRAVLLFLVLLPMLPAPPQPSGRRGRSGSGGFWGDRVDSQFPALPYIHPN 60

QY 61 PPAVDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADHTP 120
DB 61 PPAVDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPTAVAPADHTP 120

QY 121 PVPDVSRAILLRQYNLTSPVATGNTLVLYAAPLSPLLPQDGTNTTHIMATEAS 180
DB 121 PVPDVSRAILLRQYNLTSPVATGNTLVLYAAPLSPLLPQDGTNTTHIMATEAS 180

QY 181 NYAQVRVATIRYRLVFNAGVYASISFWPQTTTTSVDMNSITSDVRLVQPGI 240
DB 181 NYAQVRVATIRYRLVFNAGVYASISFWPQTTTTSVDMNSITSDVRLVQPGI 240

QY 241 ASELVTPSELHYRNQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGALGLL 300
DB 241 ASELVTPSELHYRNQWRSVETSGVAEEATSGLVMLCIHGSLVNSYNTPTGALGLL 300

QY 301 DFALELEFNLTPGNTNTRVSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFNLTPGNTNTRVSRYSSTARHLRRGADGTAEELTTAATRFMKDLYFTSTNG 360

QY 361 VGEIGRGIALTLFNLADTLGGPLTELISAGQQLFYSPVVSANGEPTVKLYTSVENAQ 420
DB 361 VGEVGRGIALTLFNLADTLGGPLTELISAGQQLFYSPVVSANGEPTVKLYTSVENAQ 420

QY 421 QDKGIAIPHIDLGESRVVIQDYDQHEQDRPTSPAPSRPFVSLRANDVWLSLTAAY 480
```

Db	421	ODKGITPHDIDLGDSRVVIQDYDNQHEQDRPTSPAPSRPFVLRANDVWLSLTAAEY	480
Qy	481	DQSTYGSSTGPFVYVSDSVTLVNVATGAQAVARSLDWTYKVTLDGRPLSTIQOYSKTFVLP	540
Db	481	DQTTYGSSTNPMYVSDTVTLVNVATGAQAVARSLDWSKVTLDGRPLTTIQOYSKTFVLP	540
Qy	541	LRGKLSFEAGTTKAGYPYNYNTTASDOLLVNNAGHRVAISTYTTSLGAGPVSISAVAV	600
Db	541	LRGKLSFEAGTTKAGYPYNYNTTASDQILIENAGHRVAISTYTTSLGAGPTSISAVGV	600
Qy	601	LAPHSALELLEDITLDYPARAHTFDDFCPECRPLGLOCAFOSTVAELQRLKMKVGTKRE	659
Db	601	LAPHSALELLEDITVDYPARAHTFDDFCPECECTLGLQCAFOSTIAELQRLKMKVGTKRE	659

Search completed: August 16, 2004, 13:35:04
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 13:33:52 ; Search time 19 seconds
(without alignments)
1793.323 Million cell updates/sec

Title: US-09-851-410A-8
Perfect score: 3408
Sequence: 1 MRPRPILLLLMLPMLPAP.....QSTVAELQRLMKVKGKTREL 660

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 segs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3408	100.0	660	1	US-08-240-049B-15
2	3408	100.0	660	1	US-08-259-148A-19
3	3408	100.0	660	1	US-08-484-054-19
4	3408	100.0	660	2	US-07-876-941A-19
5	3408	100.0	660	3	US-08-478-507-8
6	3408	100.0	660	3	US-08-542-634-13
7	3408	100.0	660	3	US-09-128-275A-8
8	3408	100.0	660	3	US-08-477-292-13
9	3408	100.0	660	4	US-09-553-427-8
10	3408	100.0	660	4	US-09-462-606-50
11	3408	100.0	660	4	US-07-870-985A-19
12	3408	100.0	660	5	PCT-US95-13703-13
13	3399	99.7	660	4	US-09-172-699-2
14	3391	99.5	660	4	US-09-462-606-12
15	3382	99.2	660	3	US-08-840-316-2
16	3382	99.2	660	3	US-08-809-523-2
17	3382	99.2	660	3	US-08-471-971-2
18	3382	99.2	660	4	US-09-462-606-49
19	3382	99.2	660	4	US-09-462-606-56
20	3382	99.2	660	4	US-09-402-776-2
21	3382	99.2	660	5	PCT-US93-08849A-2
22	3382	99.2	660	5	PCT-US93-08849-2
23	3374	99.0	660	4	US-09-462-606-51
24	3372	98.9	660	4	US-09-462-606-52
25	3367	98.8	660	4	US-09-462-606-53
26	3367	98.8	660	4	US-09-462-606-55
27	3336	97.9	660	4	US-09-462-606-54

28	3221	94.5	660	3	US-08-542-634-14	Sequence 14, Appl
29	3221	94.5	660	5	PCT-US95-13703-14	Sequence 14, Appl
30	3215	94.3	660	4	US-09-462-606-48	Sequence 48, Appl
31	3194.5	93.7	659	1	US-08-240-049B-16	Sequence 16, Appl
32	3194.5	93.7	659	1	US-08-259-148A-20	Sequence 20, Appl
33	3194.5	93.7	659	1	US-08-484-054-20	Sequence 20, Appl
34	3194.5	93.7	659	2	US-07-876-941A-20	Sequence 20, Appl
35	3194.5	93.7	659	3	US-08-477-292-14	Sequence 14, Appl
36	3194.5	93.7	659	4	US-07-870-985A-20	Sequence 20, Appl
37	3164	92.8	660	4	US-09-462-606-57	Sequence 57, Appl
38	2865	84.1	561	4	US-09-172-699-20	Sequence 20, Appl
39	2821	82.8	552	4	US-09-172-699-16	Sequence 16, Appl
40	2813	82.5	549	3	US-08-542-634-15	Sequence 15, Appl
41	2813	82.5	549	3	US-08-477-292-15	Sequence 15, Appl
42	2813	82.5	549	5	PCT-US95-13703-15	Sequence 15, Appl
43	2769	81.2	540	3	US-08-542-634-25	Sequence 25, Appl
44	2769	81.2	540	5	PCT-US95-13703-25	Sequence 25, Appl
45	2694	79.0	525	3	US-08-542-634-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-08-240-049B-15
; Sequence 15, Application US/08240049B
; Patent No. 5686239
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Tam, Albert W.
; APPLICANT: Yarbough, Patricia O.
; TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240.049B
; FILING DATE: 09-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles K. Sholtz
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
; INDIVIDUAL ISOLATE: ORF-2
US-08-240-049B-15

Query Match 100.0%; Score 3408; DB 1; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRILLLLMLPMLPAPPPQSGRRRRSGSGGFWGDRVDSQPFAPIVHTN 60

Db 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGGFWGDRVDSQFPAIPIHPTN 60
QY 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPAHDT 120
Db 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPAHDT 120
QY 121 PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLVLAAPLSPLPLQDGTNTTHIMATEAS 180
Db 121 PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLVLAAPLSPLPLQDGTNTTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYALISFWMQTTTTPTSDVMSITSDVRLVQPGI 240
Db 181 NYAQYVARATIRYRPLVNAVGGYALISFWMQTTTTPTSDVMSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVLNSYNTPTYGALGLL 300
Db 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVLNSYNTPTYGALGLL 300
QY 301 DPALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DPALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISAGAGQLFYRPPVWSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTELISAGAGQLFYRPPVWSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDGESRVVIQDYNQHEQDRPTSPAPSPFVSLRANDVLSLTAAY 480
Db 421 QDKGIAIPHDIIDGESRVVIQDYNQHEQDRPTSPAPSPFVSLRANDVLSLTAAY 480
QY 481 DQSTYSGSTGVPVYSDSVTLVNVATGAQAVARSLDWTKVLDRPLSTIOQYKTFVLP 540
Db 481 DQSTYSGSTGVPVYSDSVTLVNVATGAQAVARSLDWTKVLDRPLSTIOQYKTFVLP 540
QY 541 LRKGLSFEWAGTKAGYPYNNVTASDQLLVENAGHRVAISTVTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFEWAGTKAGYPYNNVTASDQLLVENAGHRVAISTVTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELTLDYPARAHTDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKTREL 660
Db 601 LAPHSALELTLDYPARAHTDDFCPCRLPLGQCAFQSTVAELQRLKMKVGKTREL 660

RESULT 2

US-08-259-148A-19
; Sequence 19, Application US/08259148A
; Patent No. 5741490

GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; APPLICANT: Varbough, Patrice D.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,148A
; FILING DATE: 13-JUN-1994

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA: US 505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA: US 420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA: US 367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA: US 336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA: US 208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
; US-08-259-148A-19

Query Match 100.0%; Score 3408; DB 1; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGGFWGDRVDSQFPAIPIHPTN 60
Db 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGGFWGDRVDSQFPAIPIHPTN 60
QY 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPAHDT 120
Db 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPAHDT 120
QY 121 PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLVLAAPLSPLPLQDGTNTTHIMATEAS 180
Db 121 PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLVLAAPLSPLPLQDGTNTTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVNAVGGYALISFWMQTTTTPTSDVMSITSDVRLVQPGI 240
Db 181 NYAQYVARATIRYRPLVNAVGGYALISFWMQTTTTPTSDVMSITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVLNSYNTPTYGALGLL 300
Db 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVLNSYNTPTYGALGLL 300
QY 301 DPALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DPALELEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTELISAGAGQLFYRPPVWSANGEPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTELISAGAGQLFYRPPVWSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIIDGESRVVIQDYNQHEQDRPTSPAPSPFVSLRANDVLSLTAAY 480
Db 421 QDKGIAIPHDIIDGESRVVIQDYNQHEQDRPTSPAPSPFVSLRANDVLSLTAAY 480


```
QY 481 DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARS LDWTKVTL DGRPLSTIQYQSKTFVLP 540
DB 481 DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARS LDWTKVTL DGRPLSTIQYQSKTFVLP 540
QY 541 LRGLSFWAGTTKAGYPYNYNTASDQLLVNAGHRVAISTYTTSLGAGPVSISAVAV 600
DB 541 LRGLSFWAGTTKAGYPYNYNTASDQLLVNAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSA LLEDLTLDYPARAHTFDDFCPCRPGLGQCAFSQSTVAELQRLKMKVGKTR EL 660
DB 601 LAPHSA LLEDLTLDYPARAHTFDDFCPCRPGLGQCAFSQSTVAELQRLKMKVGKTR EL 660

RESULT 3
US-08-484-054-19
; Sequence 19, Application US/08484054
; Patent No. 5770689
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,054
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 870,985
; FILING DATE: 20-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
US-08-484-054-19

Query Match 100.0%; Score 3408; DB 1; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294; Indels 0; Gaps 0;
Matches 660; Conservative 0; Mismatches 0;

QY 1 MRPRPILLLLMLFPLPAPPQPSGRRRRSGSGGFGWCDRVD SQPFAIPYIHP TN 60
DB 1 MRPRPILLLLMLFPLPAPPQPSGRRRRSGSGGFGWCDRVD SQPFAIPYIHP TN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGA PLTAVAPADTP 120
DB 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAPAVASRRRPTTAGA PLTAVAPADTP 120
QY 121 PVPDVSRCAILLRQYNLSTSLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
DB 121 PVPDVSRCAILLRQYNLSTSLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYVARATIRYRPLVPNAVGVYAISISFWPQT TTTPTSDVMNSITSDVRILVQPGI 240
DB 181 NYAQYVARATIRYRPLVPNAVGVYAISISFWPQT TTTPTSDVMNSITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYRNQGWRSVETSGVAEEETSGLVMLCIHGS LVSNTNTPTTGALGLL 300
DB 241 ASELVIPSERLHYRNQGWRSVETSGVAEEETSGLVMLCIHGS LVSNTNTPTTGALGLL 300
QY 301 DFALELEFNLTGNTNTRVSRYSSTASHRLRREGADGTAE LTTAATRFMKDLYFTSTNG 360
DB 301 DFALELEFNLTGNTNTRVSRYSSTASHRLRREGADGTAE LTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLDLTLGGLPTLGLLPGTSSAGQGFYS RPVVSANGEPVVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNLDLTLGGLPTLGLLPGTSSAGQGFYS RPVVSANGEPVVKLYTSVENAQ 420
QY 421 QDKGIAIPHDI DLGESRVVIQDYDQHEQDRTPSPAPSRPFSVLRANDV LMSLTAAEY 480
DB 421 QDKGIAIPHDI DLGESRVVIQDYDQHEQDRTPSPAPSRPFSVLRANDV LMSLTAAEY 480
QY 481 DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARS LDWTKVTL DGRPLSTIQYQSKTFVLP 540
DB 481 DQSTYGSSTGPPVYVSDSVTLNVNATGAQAVARS LDWTKVTL DGRPLSTIQYQSKTFVLP 540
QY 541 LRGLSFWAGTTKAGYPYNYNTASDQLLVNAGHRVAISTYTTSLGAGPVSISAVAV 600
DB 541 LRGLSFWAGTTKAGYPYNYNTASDQLLVNAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSA LLEDLTLDYPARAHTFDDFCPCRPGLGQCAFSQSTVAELQRLKMKVGKTR EL 660
DB 601 LAPHSA LLEDLTLDYPARAHTFDDFCPCRPGLGQCAFSQSTVAELQRLKMKVGKTR EL 660

RESULT 4
US-07-876-941A-19
; Sequence 19, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
```



```
/ FILING DATE: 17-JUN-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sholtz, Charles K.
/ REGISTRATION NUMBER: 38,615
/ REFERENCE/DOCKET NUMBER: 4600-0183.22
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 324-0880
/ TELEFAX: (650) 324-0960
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-478-507-8

Query Match      100.0%; Score 3408; DB 3; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPYIHPN 60
Db 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPYIHPN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPADTP 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPADTP 120
QY 121 PVPDVDSRGAILRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVDSRGAILRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYRPLVNAVGVYAISISFPQTTTTPTSDVMSITSTDVRIIVQPGI 240
Db 181 NYAQVRVARATIRYRPLVNAVGVYAISISFPQTTTTPTSDVMSITSTDVRIIVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSLSVNSYNTPTYGALGLL 300
Db 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSLSVNSYNTPTYGALGLL 300
QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLISAGAGQLFYSRPVVSANGEPVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLISAGAGQLFYSRPVVSANGEPVKLYTSVENAQ 420
QY 421 ODKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFVSLRANDVLMSLTAARY 480
Db 421 ODKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFVSLRANDVLMSLTAARY 480
```

RESULT 6

```
US-08-542-634-13
; Sequence 13, Application US/08542634
; Patent No. 6214970
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; APPLICANT: McAtee, C. Patrick
; APPLICANT: Yarbough, Patrice O.
```

```
/ APPLICANT: Zhang, Yifan
/ TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dehlinger & Associates
/ STREET: 350 Cambridge Ave., Suite 250
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/542,634
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0293.30
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
/ INDIVIDUAL ISOLATE: ORF-2
US-08-542-634-13
```

```
Query Match      100.0%; Score 3408; DB 3; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPYIHPN 60
Db 1 MRPRPILLLMLPMLPAPPQPSGRRGRSGSGGFGWGDVDSQPPFAIPYIHPN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPADTP 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPLTAVAPADTP 120
QY 121 PVPDVDSRGAILRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVDSRGAILRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYRPLVNAVGVYAISISFPQTTTTPTSDVMSITSTDVRIIVQPGI 240
Db 181 NYAQVRVARATIRYRPLVNAVGVYAISISFPQTTTTPTSDVMSITSTDVRIIVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSLSVNSYNTPTYGALGLL 300
Db 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSLSVNSYNTPTYGALGLL 300
QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALLEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTLISAGAGQLFYSRPVVSANGEPVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTLISAGAGQLFYSRPVVSANGEPVKLYTSVENAQ 420
QY 421 ODKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFVSLRANDVLMSLTAARY 480
Db 421 ODKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFVSLRANDVLMSLTAARY 480
```

QY 481 DQSTYGSSTGPGVYVSDSVTLVNVATGAQAVARSLDWTKVTLDRPLSTIQOYSKTFVLP 540
Db 481 DQSTYGSSTGPGVYVSDSVTLVNVATGAQAVARSLDWTKVTLDRPLSTIQOYSKTFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHRAVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHRAVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDTLDYPARAHTDDFCPCPRPLGLQCAFGSTVAELQRLKMKVGTREL 660
Db 601 LAPHSALELLEDTLDYPARAHTDDFCPCPRPLGLQCAFGSTVAELQRLKMKVGTREL 660

RESULT 7

US-09-128-275A-8
; Sequence 8, Application US/09128275A
; Patent No. 6229005
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6229005-A/No. 6229005-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,275A
; FILING DATE: 03-AUG-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-128-275A-8
Query Match 100.0%; Score 3408; DB 3; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMFLPMLPAPPPQPSGRRRGRSGGGGFWGDRVDSQFPALPYIHTN 60
Db 1 MRPRPILLLLMFLPMLPAPPPQPSGRRRGRSGGGGFWGDRVDSQFPALPYIHTN 60
QY 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAPAVASRRRPTTAGAAPTATAPAHDT 120
Db 61 PFAPDVTAAAGAGPRVRQPARPLGSAWRDQAPAVASRRRPTTAGAAPTATAPAHDT 120
QY 121 PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVSRGAILRRQYNLSTPLTSSVATGTLNLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGYAISISFWPQTTTTPTTSVDMNSTITSDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGYAISISFWPQTTTTPTTSVDMNSTITSDVRLVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSIVNSYNTPTYGALL 300
Db 241 ASELVIPSERLHYRNOGWSRVETSGVAEEATSGLVMLCIHGSIVNSYNTPTYGALL 300
QY 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELETTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELETTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLPTELISSAGQOLFYSRVPVVSANGEPYTKLYTSVNAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLPTELISSAGQOLFYSRVPVVSANGEPYTKLYTSVNAQ 420
QY 421 QDKGIAIPHDDIDGESRVVIOQYDNOHQEQRPTSPAPSRPFVLRANDVWLSTAAEY 480
Db 421 QDKGIAIPHDDIDGESRVVIOQYDNOHQEQRPTSPAPSRPFVLRANDVWLSTAAEY 480
QY 481 DQSTYGSSTGPGVYVSDSVTLVNVATGAQAVARSLDWTKVTLDRPLSTIQOYSKTFVLP 540
Db 481 DQSTYGSSTGPGVYVSDSVTLVNVATGAQAVARSLDWTKVTLDRPLSTIQOYSKTFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHRAVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAHRAVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELLEDTLDYPARAHTDDFCPCPRPLGLQCAFGSTVAELQRLKMKVGTREL 660
Db 601 LAPHSALELLEDTLDYPARAHTDDFCPCPRPLGLQCAFGSTVAELQRLKMKVGTREL 660
RESULT 8
US-08-477-292-13
; Sequence 13, Application US/08477292
; Patent No. 6291641
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas
; APPLICANT: McAtee, Patrick
; APPLICANT: Yarbough, Patrice
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cathleen M. Desjardins, M.D.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: US
; ZIP: 94063

```
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/477,292
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/327,952
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Brookes, Allan A
/ REGISTRATION NUMBER: 36,373
/ REFERENCE/DOCKET NUMBER: G32P5
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415)369-9500
/ TELEFAX: (415)368-0709
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
/ INDIVIDUAL ISOLATE: ORF-2
/ INDIVIDUAL ISOLATE:
US-08-477-292-13

Query Match      100.0%; Score 3408; DB 3; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60
Db 1 MRPRILLLLMFLPMLPAPPQSGRRRRSGSGGFWGDRVDSQPPFAIPYIHPTN 60
QY 61 PPAPDVTAAAGAPRVROPAPLGLSAWRDQAPAVASRRRTTAGAAPLTAVAPAHPTP 120
Db 61 PPAPDVTAAAGAPRVROPAPLGLSAWRDQAPAVASRRRTTAGAAPLTAVAPAHPTP 120
QY 121 PVPDVSRRGAILRRQYNLSTSLTSVATGTLNLYAAPLSPLLPLOQGTNTHIMATAS 180
Db 121 PVPDVSRRGAILRRQYNLSTSLTSVATGTLNLYAAPLSPLLPLOQGTNTHIMATAS 180
QY 181 NYAQYRVARATIRVPLVNAVGGVAISISFWPQTTTTPTSDVMSITSTDVRIILVQPGI 240
Db 181 NYAQYRVARATIRVPLVNAVGGVAISISFWPQTTTTPTSDVMSITSTDVRIILVQPGI 240
QY 241 ASELVIPSERLHYRNOGRWSVETSGVABEEATSGLVMLCIHGLSVNSYNTPTTGALGIL 300
Db 241 ASELVIPSERLHYRNOGRWSVETSGVABEEATSGLVMLCIHGLSVNSYNTPTTGALGIL 300
QY 301 DFALELEFRLNTPGNTNTRVSYSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALELEFRLNTPGNTNTRVSYSTARHRLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLLPTELISSAGGQLFYSRPVVSANGEPTVKLYTSVNAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLLPTELISSAGGQLFYSRPVVSANGEPTVKLYTSVNAQ 420
QY 421 QDKGIAIHDIDLGRSVVIQDYDNQHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
Db 421 QDKGIAIHDIDLGRSVVIQDYDNQHQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
QY 481 DQSTYGSSTGTPVYVSDSVTLNVNATGAQAVARSLDWTVTLDGRPLSTIQYKSTFFVLP 540
Db 481 DQSTYGSSTGTPVYVSDSVTLNVNATGAQAVARSLDWTVTLDGRPLSTIQYKSTFFVLP 540
QY 541 LRKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTTSLGAGPVSISAVAV 600
Db 541 LRKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTTSLGAGPVSISAVAV 600
```

RESULT 9

```
US-09-553-427-8
; Sequence 8, Application US/09553427
; Patent No. 6379891
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbrough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; HEPATITIS VIRAL AGENT
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/553,427
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0183.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-553-427-8
```

Query Match 100.0%; Score 3408; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLPMLPAPPQSGRRGRSGGGGFWGDRVDSQFPAIPYIHTPN 60
DB 1 MRPRPILLLLMLPMLPAPPQSGRRGRSGGGGFWGDRVDSQFPAIPYIHTPN 60

QY 61 PFAPDVTAAGAPRVQPARPLGSAWRDQAQPAVASRRRPTTAGAAPLTAAPAHDT 120
DB 61 PFAPDVTAAGAPRVQPARPLGSAWRDQAQPAVASRRRPTTAGAAPLTAAPAHDT 120

QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGNTHIMATEAS 180
DB 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGNTHIMATEAS 180

QY 181 NYAQYVARATIRYRPLVNAVGYAISISFWPQTPTTPTSDVMNSITSDVRLVQPGI 240
DB 181 NYAQYVARATIRYRPLVNAVGYAISISFWPQTPTTPTSDVMNSITSDVRLVQPGI 240

QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSIVNSYNTPTYGALL 300
DB 241 ASELVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSIVNSYNTPTYGALL 300

QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360

QY 361 VGEIGRGIALTLFNADTLGGLPTELISSAGQOLFYSRPVWSANGEPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNADTLGGLPTELISSAGQOLFYSRPVWSANGEPTVKLYTSVENAQ 420

QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTPSPAPSRPFSVLRANDVLMISLTAAY 480
DB 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTPSPAPSRPFSVLRANDVLMISLTAAY 480

QY 481 DQSTYSSGTPVVSVDVTLNVNATGAQAVARSLDWTKVTLDRPLSTIOQYKTFVLP 540
DB 481 DQSTYSSGTPVVSVDVTLNVNATGAQAVARSLDWTKVTLDRPLSTIOQYKTFVLP 540

QY 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAHRAVIAISTVTTSLGAGPVSISAVAV 600
DB 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAHRAVIAISTVTTSLGAGPVSISAVAV 600

QY 601 LAPHSALELDTLDYPARAHTFDDFCPCRPLGQCAFQSTVAELQRLKMKVGTREL 660
DB 601 LAPHSALELDTLDYPARAHTFDDFCPCRPLGQCAFQSTVAELQRLKMKVGTREL 660

RESULT 10

US-09-462-606-50
; Sequence 50, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; FILE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267US1
; CURRENT APPLICATION NUMBER: US/09/462,606
; CURRENT FILING DATE: 2000-06-12
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E virus
; US-09-462-606-50

Query Match 100.0%; Score 3408; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLPMLPAPPQSGRRGRSGGGGFWGDRVDSQFPAIPYIHTPN 60
DB 1 MRPRPILLLLMLPMLPAPPQSGRRGRSGGGGFWGDRVDSQFPAIPYIHTPN 60

QY 61 PFAPDVTAAGAPRVQPARPLGSAWRDQAQPAVASRRRPTTAGAAPLTAAPAHDT 120
DB 61 PFAPDVTAAGAPRVQPARPLGSAWRDQAQPAVASRRRPTTAGAAPLTAAPAHDT 120

QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGNTHIMATEAS 180
DB 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLYAAPLSPLPLQDGNTHIMATEAS 180

QY 181 NYAQYVARATIRYRPLVNAVGYAISISFWPQTPTTPTSDVMNSITSDVRLVQPGI 240
DB 181 NYAQYVARATIRYRPLVNAVGYAISISFWPQTPTTPTSDVMNSITSDVRLVQPGI 240

QY 241 ASELVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSIVNSYNTPTYGALL 300
DB 241 ASELVIPSERLHYRNOGWSRVETSGVAEBEATSGLVMLCIHGSIVNSYNTPTYGALL 300

QY 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALLEFRNLTPGNTNTRVSRYSSTARHLRGADGTAEITTTAATRFMKDLYFTSTNG 360

QY 361 VGEIGRGIALTLFNADTLGGLPTELISSAGQOLFYSRPVWSANGEPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNADTLGGLPTELISSAGQOLFYSRPVWSANGEPTVKLYTSVENAQ 420

QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTPSPAPSRPFSVLRANDVLMISLTAAY 480
DB 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTPSPAPSRPFSVLRANDVLMISLTAAY 480

QY 481 DQSTYSSGTPVVSVDVTLNVNATGAQAVARSLDWTKVTLDRPLSTIOQYKTFVLP 540
DB 481 DQSTYSSGTPVVSVDVTLNVNATGAQAVARSLDWTKVTLDRPLSTIOQYKTFVLP 540

QY 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAHRAVIAISTVTTSLGAGPVSISAVAV 600
DB 541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAHRAVIAISTVTTSLGAGPVSISAVAV 600

QY 601 LAPHSALELDTLDYPARAHTFDDFCPCRPLGQCAFQSTVAELQRLKMKVGTREL 660
DB 601 LAPHSALELDTLDYPARAHTFDDFCPCRPLGQCAFQSTVAELQRLKMKVGTREL 660

RESULT 11

US-07-870-985A-19
; Sequence 19, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Ji-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/870,985A
/ FILING DATE: 20-APRIL-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 822,335
/ FILING DATE: 17-JAN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 505,888
/ FILING DATE: 05-APRIL-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 420,921
/ FILING DATE: 13-OCTOBER-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 367,486
/ FILING DATE: 16-JUNE-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 336,672
/ FILING DATE: 11-APRIL-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 208,997
/ FILING DATE: 17-JUNE-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0093.30
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
/ US-07-870-985A-19

Query Match 100.0%; Score 3408; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRILLLLMLPAPPQSGRRRRSGSGGFGWDRVDSQPPAIPYIHPTN 60
Db 1 MRPRILLLLMLPAPPQSGRRRRSGSGGFGWDRVDSQPPAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQORPAVASRRRTTAGAAPLTAAPAHDTP 120
Db 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQORPAVASRRRTTAGAAPLTAAPAHDTP 120
QY 121 FVPDVSRRGAILRRQYNLSPTLSVATGNNLVLYAAPLSPLLPQDGTNTHIMATEAS 180
Db 121 FVPDVSRRGAILRRQYNLSPTLSVATGNNLVLYAAPLSPLLPQDGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRPLVPNAVGGVAISIFWPQTTPTSVDMNSITSDVRILVQPGI 240
Db 181 NYAQVRVARATIRPLVPNAVGGVAISIFWPQTTPTSVDMNSITSDVRILVQPGI 240
QY 241 ASELVIPSRLHYRQGRSVETSGVAEEATSGLVMLCIHGLSVNSYNTPYTCALGLL 300
Db 241 ASELVIPSRLHYRQGRSVETSGVAEEATSGLVMLCIHGLSVNSYNTPYTCALGLL 300
QY 301 DFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTTAATRFMKDLYFTSTNG 360
Db 301 DFALEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAELTTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLGLPTLTISSAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLGLPTLTISSAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420

/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/870,985A
/ FILING DATE: 20-APRIL-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 822,335
/ FILING DATE: 17-JAN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 505,888
/ FILING DATE: 05-APRIL-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 420,921
/ FILING DATE: 13-OCTOBER-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 367,486
/ FILING DATE: 16-JUNE-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 336,672
/ FILING DATE: 11-APRIL-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 208,997
/ FILING DATE: 17-JUNE-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0093.30
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/13703
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0293.41
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
/ INDIVIDUAL ISOLATE: ORF-2
/ PCT-US95-13703-13

Query Match 100.0%; Score 3408; DB 5; Length 660;
Best Local Similarity 100.0%; Pred. No. 3.7e-294;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRILLLLMLPAPPQSGRRRRSGSGGFGWDRVDSQPPAIPYIHPTN 60
Db 1 MRPRILLLLMLPAPPQSGRRRRSGSGGFGWDRVDSQPPAIPYIHPTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQORPAVASRRRTTAGAAPLTAAPAHDTP 120
Db 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQORPAVASRRRTTAGAAPLTAAPAHDTP 120

/ RESULT 12
/ PCT-US95-13703-13
/ Sequence 13, Application PC/TUS9513703
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
/ TITLE OF INVENTION: USES THEREFOR
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dehlinger & Associates
/ STREET: P.O. Box 60850
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306-0850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/13703
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0293.41
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
/ INDIVIDUAL ISOLATE: ORF-2
/ PCT-US95-13703-13
```

QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLVYAAPLSPLLPQDGTNTHIMATEAS 180
DB 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLVYAAPLSPLLPQDGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYRPLVNAVGGYVAISISFWPQTITTTTSDVMSITSDVRLVQPGI 240
DB 181 NYAQVRVARATIRYRPLVNAVGGYVAISISFWPQTITTTTSDVMSITSDVRLVQPGI 240
QY 241 ASELVTPSRHLHYRNOGWSVETSGVAEBEATSGLVMLCIHGLSVNSYNTPTGTALGLL 300
DB 241 ASELVTPSRHLHYRNOGWSVETSGVAEBEATSGLVMLCIHGLSVNSYNTPTGTALGLL 300
QY 301 DPALELEFNLTPGNTNTRVSRYSSTARHRLRGADGTAELETTAATRFMKDLYFTSTNG 360
DB 301 DPALELEFNLTPGNTNTRVSRYSSTARHRLRGADGTAELETTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNLADTLLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFVLRANDVWLMLSLTAAEY 480
DB 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFVLRANDVWLMLSLTAAEY 480
QY 481 DOSTYSSSTGPVVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTTQQYSKTFVFLP 540
DB 481 DOSTYSSSTGPVVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTTQQYSKTFVFLP 540
QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTTSLGAGPVSISAVAV 600
DB 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTTSLGAGPVSISAVAV 600
QY 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPGLGQCAFOSTVAELQRLKMKVGTREL 660
DB 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPGLGQCAFOSTVAELQRLKMKVGTREL 660

RESULT 13

US-09-172-699-2

; Sequence 2, Application US/09172699A
; Patent No. 6514690
; GENERAL INFORMATION:
; APPLICANT: Anderson, David A.
; APPLICANT: Locarnini, Stephen A.
; APPLICANT: Toressi, Joseph
; APPLICANT: Hui, Zhuang
; APPLICANT: Li, Fan
; TITLE OF INVENTION: IMMUNOREACTIVE ANTIGENS OF HEPATITIS E VIRUS
; FILE REFERENCE: Davies Col. Cave
; CURRENT APPLICATION NUMBER: US/09/172,699A
; CURRENT FILING DATE: 1998-10-14
; EARLIER APPLICATION NUMBER: 08/617,927
; EARLIER FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-172-699-2

Query Match 99.78; Score 3399; DB 4; Length 660;
Best Local Similarity 99.78; Pred. No. 2.3e-293;
Matches 658; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGFWGDRVDSQFAIPIYHPTN 60
DB 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGFWGDRVDSQFAIPIYHPTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQRPVAVASRRRPTTAGAAPTAVAPADHTP 120
DB 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQRPVAVASRRRPTTAGAAPTAVAPADHTP 120

QY 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLVYAAPLSPLLPQDGTNTHIMATEAS 180
DB 121 PVPDVSRAILRRQYNLSPLTSSVATGTLNLVYAAPLSPLLPQDGTNTHIMATEAS 180
QY 181 NYAQVRVARATIRYRPLVNAVGGYVAISISFWPQTITTTTSDVMSITSDVRLVQPGI 240
DB 181 NYAQVRVARATIRYRPLVNAVGGYVAISISFWPQTITTTTSDVMSITSDVRLVQPGI 240
QY 241 ASELVTPSRHLHYRNOGWSVETSGVAEBEATSGLVMLCIHGLSVNSYNTPTGTALGLL 300
DB 241 ASELVTPSRHLHYRNOGWSVETSGVAEBEATSGLVMLCIHGLSVNSYNTPTGTALGLL 300
QY 301 DPALELEFNLTPGNTNTRVSRYSSTARHRLRGADGTAELETTAATRFMKDLYFTSTNG 360
DB 301 DPALELEFNLTPGNTNTRVSRYSSTARHRLRGADGTAELETTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNLADTLLGGLPTELISSAGQOLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFVLRANDVWLMLSLTAAEY 480
DB 421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFVLRANDVWLMLSLTAAEY 480
QY 481 DOSTYSSSTGPVVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTTQQYSKTFVFLP 540
DB 481 DOSTYSSSTGPVVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTTQQYSKTFVFLP 540
QY 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTTSLGAGPVSISAVAV 600
DB 541 LRKGLSFEWAGTTKAGYPYNYNTTASDQLLVENAGHRVAISTYTTTSLGAGPVSISAVAV 600
QY 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPGLGQCAFOSTVAELQRLKMKVGTREL 660
DB 601 LAPHSAALLEDTLDYPARAHTFDDFCPCRPGLGQCAFOSTVAELQRLKMKVGTREL 660

RESULT 14

US-09-462-606-12

; Sequence 12, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267US1
; CURRENT APPLICATION NUMBER: US/09/462,606
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-12

Query Match 99.5%; Score 3391; DB 4; Length 660;
Best Local Similarity 99.4%; Pred. No. 1.2e-292;
Matches 656; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGFWGDRVDSQFAIPIYHPTN 60
DB 1 MRPRPILLLLMLPMLPAPPQPSGRRGRSGSGGFWGDRVDSQFAIPIYHPTN 60
QY 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQRPVAVASRRRPTTAGAAPTAVAPADHTP 120
DB 61 PFAPDVTAAAGAGPRVQPARPLGSAWRDQAQRPVAVASRRRPTTAGAAPTAVAPADHTP 120


```
QY 121 PVPDVDSRGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVDSRGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYPLVPNAVGGVAISISFWPQTITTTTPTSDVMSITSTDVRIILVQPGI 240
Db 181 NYAQYRVARATIRYPLVPNAVGGVAISISFWPQTITTTTPTSDVMSITSTDVRIILVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVLNSVNTNTPTGTALGLL 300
Db 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVLNSVNTNTPTGTALGLL 300
QY 301 DFALBEPFNLTGNTNTRVSYSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
Db 301 DFALBEPFNLTGNTNTRVSYSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTLTISSAGQLFYSPVPVSANGPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTLTISSAGQLFYSPVPVSANGPTVKLYTSVENAQ 420
QY 421 QDKGTAIPHDIIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVWLSLTAABY 480
Db 421 QDKGTAIPHDIIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVWLSLTAABY 480
QY 481 DOSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTTQQYSKTFFVLP 540
Db 481 DOSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTTQQYSKTFFVLP 540
QY 541 LRGKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRGKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELTDLDYPARAHTFDDPCPCPLGLQGCAFSQSTVAELQRLKMKVGKTRIL 660
Db 601 LAPHSALELTDLDYPARAHTFDDPCPCPLGLQGCAFSQSTVAELQRLKMKVGKTRIL 660
```

RESULT 15

```
US-08-840-316-2
; Sequence 2, Application US/08840316
; Patent No. 6054567
; GENERAL INFORMATION:
; APPLICANT: Emerson, Suzanne U., Purcell, Robert H.,
; APPLICANT: Tearev, Sergei. A., and Robinson, Robin A.
; TITLE OF INVENTION: Recombinant Proteins Of
; TITLE OF INVENTION: A Pakistani Strain Of Hepatitis E And Their
; TITLE OF INVENTION: Use In Diagnostic Methods And Vaccines
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,316
; FILING DATE: 11-APR-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard W. Bork
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4255
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
US-08-840-316-2
```

```
Query Match 99.2%; Score 3382; DB 3; Length 660;
Best Local Similarity 99.2%; Pred. No. 7.6e-292;
Matches 655; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 MRPRLTLLLLMFLPMLPAPPQPSGRRGRSGSGGFWGDRVDSOPEALPIYHPTN 60
Db 1 MRPRLTLLLLMFLPMLPAPPQPSGRRGRSGSGGFWGDRVDSOPEALPIYHPTN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSARWDOQRPVAVASRRRPTTAGAAPLTAFAHADTP 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSARWDOQRPVAVASRRRPTTAGAAPLTAFAHADTP 120
QY 121 PVPDVDSRGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVDSRGAILRRQYNLSTSLTSSVATGNTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYPLVPNAVGGVAISISFWPQTITTTTPTSDVMSITSTDVRIILVQPGI 240
Db 181 NYAQYRVARATIRYPLVPNAVGGVAISISFWPQTITTTTPTSDVMSITSTDVRIILVQPGI 240
QY 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVLNSVNTNTPTGTALGLL 300
Db 241 ASELVIPSERLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVLNSVNTNTPTGTALGLL 300
QY 301 DFALBEPFNLTGNTNTRVSYSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
Db 301 DFALBEPFNLTGNTNTRVSYSTARHRLRGADGTAEELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGGLPTLTISSAGQLFYSPVPVSANGPTVKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGGLPTLTISSAGQLFYSPVPVSANGPTVKLYTSVENAQ 420
QY 421 QDKGTAIPHDIIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVWLSLTAABY 480
Db 421 QDKGTAIPHDIIDLGESRVVIQDYDNQHEQDRPTSPAPSRPFSVLRANDVWLSLTAABY 480
QY 481 DOSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTTQQYSKTFFVLP 540
Db 481 DOSTYGSSTGPPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDCRPLSTTQQYSKTFFVLP 540
QY 541 LRGKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRGKLSFWEAGTTKAGYPYNYNTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHSALELTDLDYPARAHTFDDPCPCPLGLQGCAFSQSTVAELQRLKMKVGKTRIL 660
Db 601 LAPHSALELTDLDYPARAHTFDDPCPCPLGLQGCAFSQSTVAELQRLKMKVGKTRIL 660
```

Search completed: August 16, 2004, 13:36:07
Job time : 21 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2004, 13:35:07 ; Search time 51 Seconds
(without alignments)
4062.587 Million cell updates/sec

Title: US-09-851-410A-8
Perfect score: 3408
Sequence: 1 MRPRILLLLMFLPMLPAP.....QSTVAELQRLKMKVGKTRRL 660

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	3408	100.0	660	9	US-09-769-066-13
2	3408	100.0	660	10	US-09-851-410-8
3	3408	100.0	660	14	US-10-165-868-19
4	3391	99.5	660	12	US-10-381-770-1
5	3383	99.3	660	15	US-10-257-044-1
6	3321	94.5	660	9	US-09-769-066-14
7	3194.5	93.7	659	14	US-10-165-868-20
8	3179	93.3	674	16	US-10-239-090A-57
9	3176	93.2	660	16	US-10-239-090A-51
10	3174	93.1	660	16	US-10-239-090A-53
11	3166	92.9	674	16	US-10-239-090A-55
12	3160	92.7	660	10	US-09-468-147-92
13	3160	92.7	660	12	US-10-319-745-92
14	3151	92.5	660	10	US-09-468-147-167
15	3151	92.5	660	12	US-10-319-745-167

16	2813	82.5	549	9	US-09-769-066-15	Sequence 15, Appl
17	2769	79.0	540	9	US-09-769-066-25	Sequence 25, Appl
18	2694	79.0	525	9	US-09-769-066-27	Sequence 27, Appl
19	2684	78.8	549	9	US-09-769-066-16	Sequence 16, Appl
20	2644	77.6	540	9	US-09-769-066-26	Sequence 26, Appl
21	2569	75.4	525	9	US-09-769-066-28	Sequence 28, Appl
22	2232	65.5	436	14	US-10-165-868-17	Sequence 17, Appl
23	2098.5	61.6	435	14	US-10-165-868-18	Sequence 18, Appl
24	1677	49.2	327	9	US-09-769-066-17	Sequence 17, Appl
25	1677	49.2	327	14	US-10-165-868-15	Sequence 15, Appl
26	1615	47.4	459	10	US-09-468-147-206	Sequence 206, App
27	1615	47.4	459	10	US-09-468-147-207	Sequence 207, App
28	1615	47.4	459	12	US-10-319-745-206	Sequence 206, App
29	1615	47.4	459	12	US-10-319-745-207	Sequence 207, App
30	1601	47.0	338	10	US-09-468-147-199	Sequence 199, App
31	1601	47.0	338	10	US-09-468-147-200	Sequence 200, App
32	1601	47.0	338	12	US-10-319-745-199	Sequence 199, App
33	1601	47.0	338	12	US-10-319-745-200	Sequence 200, App
34	1594	46.8	327	9	US-09-769-066-18	Sequence 18, Appl
35	1594	46.8	327	14	US-10-165-868-16	Sequence 16, Appl
36	1592	46.7	327	10	US-09-468-147-176	Sequence 176, App
37	1592	46.7	327	12	US-10-319-745-176	Sequence 176, App
38	1581	46.4	327	10	US-09-468-147-175	Sequence 175, App
39	1581	46.4	327	12	US-10-319-745-175	Sequence 175, App
40	1352	39.7	276	10	US-09-468-147-189	Sequence 189, App
41	1352	39.7	276	12	US-10-319-745-189	Sequence 189, App
42	1334.5	39.2	606	14	US-10-029-840-6	Sequence 6, Appli
43	1273	37.4	389	12	US-10-381-770-4	Sequence 4, Appli
44	1084	31.8	232	12	US-10-381-770-12	Sequence 12, Appl
45	1076	31.6	214	12	US-10-381-770-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-769-066-13
; Sequence 13, Application US/09769066
; Patent No. US20020107360A1

GENERAL INFORMATION:

APPLICANT: Fuerst, Thomas R.
McAtee, C. Patrick
Yarbough, Patrice O.
Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates
CITY: Palo Alto
STATE: CA
COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/769,066
FILING DATE: 24-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/542,634
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33 875
REFERENCE/DOCKET NUMBER: 4600-0293.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 13:

```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 660 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: Hepatitis E Virus (Burma strain)
;   ORF-2
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
;   ORIGINAL SOURCE:
;   SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-769-066-13

Query Match          100.0%; Score 3408; DB 9; Length 660;
Best Local Similarity 100.0%; Pred. No. 7.9e-276;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLFPLPAPPPQPSGRRRGRSGSGGGFGWGDVDSQFFAIPYIHPTN 60
DB 1 MRPRPILLLLMLFPLPAPPPQPSGRRRGRSGSGGGFGWGDVDSQFFAIPYIHPTN 60

QY 61 PFAPDVTAAGAGPRVROPARPLGSARWDOAQPASVARRRPTTAGAAPTAVAPAHDT 120
DB 61 PFAPDVTAAGAGPRVROPARPLGSARWDOAQPASVARRRPTTAGAAPTAVAPAHDT 120

QY 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGNTNLVLYAAPLSPLLPLODGTNTHIMATEAS 180
DB 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGNTNLVLYAAPLSPLLPLODGTNTHIMATEAS 180

QY 181 NYAQYRVARATIRYRPLVPNAVGGYALISFQWQTTTTPTSDVMNSITSDVRLVQPGI 240
DB 181 NYAQYRVARATIRYRPLVPNAVGGYALISFQWQTTTTPTSDVMNSITSDVRLVQPGI 240

QY 241 ASLVIPSERLHYRNOGWRVSVETSGVAEEATSGLVMLCIHGSIVNSYTNPTYTGAIGLL 300
DB 241 ASLVIPSERLHYRNOGWRVSVETSGVAEEATSGLVMLCIHGSIVNSYTNPTYTGAIGLL 300

QY 301 DFALAEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
DB 301 DFALAEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360

QY 361 VGEIGRGIALTLFNLADTLGLLPTELISAGGGLFYRSPVWSANGEPYTKLYTSVENAQ 420
DB 361 VGEIGRGIALTLFNLADTLGLLPTELISAGGGLFYRSPVWSANGEPYTKLYTSVENAQ 420

QY 421 QDKGIAIPHDIIDGESRVVIQDYDQHEQDRPTSPAPSPRPFVLRANDVWLSTAAEY 480
DB 421 QDKGIAIPHDIIDGESRVVIQDYDQHEQDRPTSPAPSPRPFVLRANDVWLSTAAEY 480

QY 481 DQSTYGSSTGPVVVSDSVTLNVNATGAQAVARSLDWTKVTLDRPLSTIOQYKSTFFVLP 540
DB 481 DQSTYGSSTGPVVVSDSVTLNVNATGAQAVARSLDWTKVTLDRPLSTIOQYKSTFFVLP 540

QY 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRAVIASTYTTSLGAGPVSISAVV 600
DB 541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRAVIASTYTTSLGAGPVSISAVV 600

QY 601 LAPHSALEEDLDYPARAHTDDCPCEKPLGLQCAQSTVAELQRLKMKVGTREL 660
DB 601 LAPHSALEEDLDYPARAHTDDCPCEKPLGLQCAQSTVAELQRLKMKVGTREL 660
```

RESULT 2

```
US-09-851-410-8
; Sequence 8, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yarbough, Patrice O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; Fry, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
```

```
;
; No. US20030124510A1-A/No. US20030124510A1-B Hepatitis Viru
;
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,410
; FILING DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 660 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-851-410-8
```

```
Query Match          100.0%; Score 3408; DB 10; Length 660;
Best Local Similarity 100.0%; Pred. No. 7.9e-276;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRPRPILLLLMLFPLPAPPPQPSGRRRGRSGSGGGFGWGDVDSQFFAIPYIHPTN 60
DB 1 MRPRPILLLLMLFPLPAPPPQPSGRRRGRSGSGGGFGWGDVDSQFFAIPYIHPTN 60

QY 61 PFAPDVTAAGAGPRVROPARPLGSARWDOAQPASVARRRPTTAGAAPTAVAPAHDT 120
DB 61 PFAPDVTAAGAGPRVROPARPLGSARWDOAQPASVARRRPTTAGAAPTAVAPAHDT 120

QY 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGNTNLVLYAAPLSPLLPLODGTNTHIMATEAS 180
DB 121 PVPDVDSRGAILRRQYNLSTPLTSSVATGNTNLVLYAAPLSPLLPLODGTNTHIMATEAS 180

QY 181 NYAQYRVARATIRYRPLVPNAVGGYALISFQWQTTTTPTSDVMNSITSDVRLVQPGI 240
DB 181 NYAQYRVARATIRYRPLVPNAVGGYALISFQWQTTTTPTSDVMNSITSDVRLVQPGI 240

QY 241 ASLVIPSERLHYRNOGWRVSVETSGVAEEATSGLVMLCIHGSIVNSYTNPTYTGAIGLL 300
DB 241 ASLVIPSERLHYRNOGWRVSVETSGVAEEATSGLVMLCIHGSIVNSYTNPTYTGAIGLL 300

QY 301 DFALAEFRNLTPGNTNTRVSRYSSTARHRLRGADGTAEITTTAATRFMKDLYFTSTNG 360
```

Db 301 DPALEFRNLTPGNTNTRVSYSTARHRRGADGTAEELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLPLTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLPLTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGSRVVIQDYDNQHEODRTPSPAPSRPFSVLRANDVLMSLTAAEY 480
Db 421 QDKGIAIPHDIDLGSRVVIQDYDNQHEODRTPSPAPSRPFSVLRANDVLMSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQQYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQQYSKTFFVLP 540
QY 541 LRKGLSFWAAGTKAGYPYNNYNTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWAAGTKAGYPYNNYNTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHASALALLEDTLDYPARAHTFDDPCPCRLPLGQCAFQSTVAELQRLKMKVGKTREL 660
Db 601 LAPHASALALLEDTLDYPARAHTFDDPCPCRLPLGQCAFQSTVAELQRLKMKVGKTREL 660

RESULT 3

US-10-165-868-19
; Sequence 19, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Bradley, Daniel W.
; Twu, Jr-Shin
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; FILING DATE: 06-June-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-165-868-19

Query Match 100.0%; Score 3408; DB 14; Length 660;
Best Local Similarity 100.0%; Pred. No. 7.9e-276; Indels 0; Gaps 0;
Matches 660; Conservative 0; Mismatches 0;

QY 1 MRPRPILLLLMFLPMLPAPPQPSGRRRGRSGSGGFWGDRVDSQFFFAIPYIHPN 60
Db 1 MRPRPILLLLMFLPMLPAPPQPSGRRRGRSGSGGFWGDRVDSQFFFAIPYIHPN 60
QY 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPTAVAPADTP 120
Db 61 PFAPDVTAAGAGPRVQPARPLGSAWRDQAORPAVASRRRPTTAGAAPTAVAPADTP 120
QY 121 PVPDVDSGAILRRQYNLSTSLTSSVATGTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
Db 121 PVPDVDSGAILRRQYNLSTSLTSSVATGTLVLYAAPLSPLLPLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVENVAGGYAISISFPQPTTTTPTSDVMSITSTDVRLVQPGI 240
Db 181 NYAQYRVARATIRYRPLVENVAGGYAISISFPQPTTTTPTSDVMSITSTDVRLVQPGI 240
QY 241 ASLVIPERLHYRQNGWRSVETSGVAEBEATSGLVMLCIHGSLSVNSYNTPTTGALGL 300
Db 241 ASLVIPERLHYRQNGWRSVETSGVAEBEATSGLVMLCIHGSLSVNSYNTPTTGALGL 300
QY 301 DPALEFRNLTPGNTNTRVSYSTARHRRGADGTAEELTTAATRFMKDLYFTSTNG 360
Db 301 DPALEFRNLTPGNTNTRVSYSTARHRRGADGTAEELTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNLADTLGLPLTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLPLTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420
QY 421 QDKGIAIPHDIDLGSRVVIQDYDNQHEODRTPSPAPSRPFSVLRANDVLMSLTAAEY 480
Db 421 QDKGIAIPHDIDLGSRVVIQDYDNQHEODRTPSPAPSRPFSVLRANDVLMSLTAAEY 480
QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQQYSKTFFVLP 540
Db 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQQYSKTFFVLP 540
QY 541 LRKGLSFWAAGTKAGYPYNNYNTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRKGLSFWAAGTKAGYPYNNYNTASDQLLVENAGHRVAISTYTTSLGAGPVSISAVAV 600
QY 601 LAPHASALALLEDTLDYPARAHTFDDPCPCRLPLGQCAFQSTVAELQRLKMKVGKTREL 660
Db 601 LAPHASALALLEDTLDYPARAHTFDDPCPCRLPLGQCAFQSTVAELQRLKMKVGKTREL 660

RESULT 4
US-10-381-770-1
; Sequence 1, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS
; FILE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
; FILE REFERENCE: IEC010037PCT
; CURRENT APPLICATION NUMBER: US/10/381,770

; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 660
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-1

Query Match 99.5%; Score 3391; DB 12; Length 660;
Best Local Similarity 99.4%; Pred. No. 2.1e-274;
Matches 656; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY	1	MRPRPILLLMLPMLPAPPQPSGRRRRSGSGGFGWGDVDSQPPAIPYIHTN	60
DB	1	MRPRPILLLMLPMLPAPPQPSGRRRRSGSGGFGWGDVDSQPPAIPYIHTN	60
QY	61	PFAPDVTAAAGAGPRVQPARPLGSARWDOAQRPAVASRRRPTTAGAAPTAVAPADHTP	120
DB	61	PFAPDVTAAAGAGPRVQPARPLGSARWDOAQRPAVASRRRPTTAGAAPTAVAPADHTP	120
QY	121	PVPDVSRGAILRRQYNLSTPLTSSVATGTLNVLVYAAPLSPLLPQDGTNTHIMATEAS	180
DB	121	PVPDVSRGAILRRQYNLSTPLTSSVATGTLNVLVYAAPLSPLLPQDGTNTHIMATEAS	180
QY	181	NYAQYRVARATIRYRPLVNAVGVYAISISFWPQTTTTPTSDVMNSITSTDVRLVQPGI	240
DB	181	NYAQYRVARATIRYRPLVNAVGVYAISISFWPQTTTTPTSDVMNSITSTDVRLVQPGI	240
QY	241	ASELVIPSERLHYRNQWRSVETSGVAEEBATSGLVMLCIHGSVNSYTNTPYTGAJGLL	300
DB	241	ASELVIPSERLHYRNQWRSVETSGVAEEBATSGLVMLCIHGSVNSYTNTPYTGAJGLL	300
QY	301	DFALELFRNLTPGNTNTRVSRYSSTARHLRGGADGTAEITTTAATFRMKDLYFTSTNG	360
DB	301	DFALELFRNLTPGNTNTRVSRYSSTARHLRGGADGTAEITTTAATFRMKDLYFTSTNG	360
QY	361	VGEIGRGIALTLFNADTLGGPTELISSAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ	420
DB	361	VGEIGRGIALTLFNADTLGGPTELISSAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ	420
QY	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSPFVLRANDVWLSTAAEY	480
DB	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSPFVLRANDVWLSTAAEY	480
QY	481	DQSTYGSSTGVPVVSVDVTLVNVATGAQAVARSLDWTKVLDRPLSTIQOYSKTFFVLP	540
DB	481	DQSTYGSSTGVPVVSVDVTLVNVATGAQAVARSLDWTKVLDRPLSTIQOYSKTFFVLP	540
QY	541	LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTVTTSLGAGPVSISAVV	600
DB	541	LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTVTTSLGAGPVSISAVV	600
QY	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPLGLQCAFSQSTVAELQRLKMKVGKTREL	660
DB	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPLGLQCAFSQSTVAELQRLKMKVGKTREL	660

RESULT 5

US-10-257-044-1
; Sequence 1, Application US/10257044
; Publication No. US20030220475A1
; GENERAL INFORMATION:

; APPLICANT: The Government of the United States of America, as
; APPLICANT: Represented by the Secretary, Department of Health and
; APPLICANT: Human Services, Centers for Disease Control and Prevention
; APPLICANT: HOWARD A. FIELDS
; APPLICANT: YURI E. KHUDYAKOV
; APPLICANT: JIHONG MENG
; TITLE OF INVENTION: NEUTRALIZING IMMUNOGENIC HEV POLYPEPTIDES (HEV)
; FILE REFERENCE: 14114.034002
; CURRENT APPLICATION NUMBER: US/10/257,044
; CURRENT FILING DATE: 2002-10-04

; PRIOR APPLICATION NUMBER: 60/195,380
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:/ No. US20030220475ale =

; OTHER INFORMATION: Synthetic construct
US-10-257-044-1

Query Match 99.3%; Score 3383; DB 15; Length 660;

Best Local Similarity 99.4%; Pred. No. 9.7e-214;
Matches 656; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY	1	MRPRPILLLMLPMLPAPPQPSGRRRRSGSGGFGWGDVDSQPPAIPYIHTN	60
DB	1	MRPRPILLLMLPMLPAPPQPSGRRRRSGSGGFGWGDVDSQPPAIPYIHTN	60
QY	61	PFAPDVTAAAGAGPRVQPARPLGSARWDOAQRPAVASRRRPTTAGAAPTAVAPADHTP	120
DB	61	PFAPDVTAAAGAGPRVQPARPLGSARWDOAQRPAVASRRRPTTAGAAPTAVAPADHTP	120
QY	121	PVPDVSRGAILRRQYNLSTPLTSSVATGTLNVLVYAAPLSPLLPQDGTNTHIMATEAS	180
DB	121	PVPDVSRGAILRRQYNLSTPLTSSVATGTLNVLVYAAPLSPLLPQDGTNTHIMATEAS	180
QY	181	NYAQYRVARATIRYRPLVNAVGVYAISISFWPQTTTTPTSDVMNSITSTDVRLVQPGI	240
DB	181	NYAQYRVARATIRYRPLVNAVGVYAISISFWPQTTTTPTSDVMNSITSTDVRLVQPGI	240
QY	241	ASELVIPSERLHYRNQWRSVETSGVAEEBATSGLVMLCIHGSVNSYTNTPYTGAJGLL	300
DB	241	ASELVIPSERLHYRNQWRSVETSGVAEEBATSGLVMLCIHGSVNSYTNTPYTGAJGLL	300
QY	301	DFALELFRNLTPGNTNTRVSRYSSTARHLRGGADGTAEITTTAATFRMKDLYFTSTNG	360
DB	301	DFALELFRNLTPGNTNTRVSRYSSTARHLRGGADGTAEITTTAATFRMKDLYFTSTNG	360
QY	361	VGEIGRGIALTLFNADTLGGPTELISSAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ	420
DB	361	VGEIGRGIALTLFNADTLGGPTELISSAGGOLFYSRPVVSANGEPPTVKLYTSVENAQ	420
QY	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSPFVLRANDVWLSTAAEY	480
DB	421	QDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTSPAPSPFVLRANDVWLSTAAEY	480
QY	481	DQSTYGSSTGVPVVSVDVTLVNVATGAQAVARSLDWTKVLDRPLSTIQOYSKTFFVLP	540
DB	481	DQSTYGSSTGVPVVSVDVTLVNVATGAQAVARSLDWTKVLDRPLSTIQOYSKTFFVLP	540
QY	541	LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTVTTSLGAGPVSISAVV	600
DB	541	LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTVTTSLGAGPVSISAVV	600
QY	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPLGLQCAFSQSTVAELQRLKMKVGKTREL	660
DB	601	LAPHSALALLEDTLDYPARAHTFDDFCPCRPLGLQCAFSQSTVAELQRLKMKVGKTREL	660

RESULT 6

US-09-769-066-14
; Sequence 14, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:

; APPLICANT: Fuerst, Thomas R.
; McAttee, C. Patrick
; Yarbrough, Patrice O.
; Zhang, Yifan

; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31

```
/
/
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Dehlinger & Associates
/ STREET: 350 Cambridge Ave., Suite 250
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/769,066
/ FILING DATE: 24-Jan-2001
/ CLASSIFICATION: <Unknown>
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/542,634
/ FILING DATE: <Unknown>
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 4600-0293.30
/ REFERENCE/DOCKET NUMBER: 4600-0293.30
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 660 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: Hepatitis E Virus (Mexico Strain)
/ ORF-2
/
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/
/ ORIGINAL SOURCE:
/ SEQUENCE DESCRIPTION: SEQ ID NO: 14:
/
/ US-09-769-066-14
/
/
/ Query Match 94.5%; Score 3221; DB 9; Length 660;
/ Best Local Similarity 93.5%; Pred. No. 3.4e-260;
/ Matches 617; Conservative 21; Mismatches 22; Indels 0; Gaps 0;
/
/ QY 1 MRPRPILLLMLPMPAPPQPSGRRRRSGSGGFGWDRVDSQPFAPYTHPTN 60
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 1 MRPRPILLLFLPLMPAPPTQPSGRRRRSGSGTGGFGWDRVDSQPFAPYTHPTN 60
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 61 PFAPDVTAAGAPRVQPARPLGSAWRDQAPAVASRRRPTTAGAAPLTAVAPADHTP 120
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 61 PFAPDVAAASGSGPRLRQPARPLGSGTWDRDQAPSAASRRRPTAGAAALTAVAPADHTS 120
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 121 PVPDVSRCAILRROYNLSTSLTSSVATGTLNLYAAPLSPLPLOGDGTNTHIMATEAS 180
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 121 PVPDVSRCAILRQYNLSTSLTSSVAGTGLNLYAAPLPPLQDGTNTHIMATEAS 180
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 181 NYAQVRVATIRYRPLVNAVGGVAISISFWPQTITTTTSVDMNSITSDVRLVQPGI 240
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 181 NYAQVRVATIRYRPLVNAVGGVAISISFWPQTITTTTSVDMNSITSDVRLVQPGI 240
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 241 ASELVPSRLHYRNOGWSVETSGVAEEATSGLVMLCIHGSVLNSYNTPTYGALGLL 300
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 301 DPALELEFNLTPGNTNTRVSYSSSTARHLRGGADGTAELTTAATRFMKDLYFTSTNG 360
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 361 DPALELEFNLTTCNTNTRVSYSSSTARHLRGGADGTAELTTAATRFMKDLHFTGLNG 360
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 361 VGEIGRGIALTLFNLADTLGLLPTELISAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 361 VGEVGRGIALTLNLADTLGLLPTELISAGGQLFYSRPVVSANGEPVTKLYTSVENAQ 420
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 421 QDKGIAIPHDIDIGESRVVIQDYDNQHQEDRTPSPAPSRPFSVLRANDVLMLSLTAABY 480
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/
/
/ 421 QDKGVAIPHDIDIGDSRVVIQDYDNQHQEDRTPSPAPSRPFSVLRANDVLMLSLTAABY 480
/ QY 481 DOSTYGSSTGPPYVSDSVTLNVNATCAQAVARSLDWTKVTLQGRPLSTTQQYSKTFVLP 540
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 481 DOSTYGSSTGPPYVSDSVTLNVNATCAQAVARSLDWTKVTLQGRPLSTTQQYSKTFVLP 540
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 541 LRKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLGAGPVYSISAVAV 600
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 541 LRKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLGAGPVYSISAVAV 600
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 601 LAPHSALALLEDTLDYPARAHTFDDFCPCRPGLGQGCFAQFQSTVABLQRLKMKVGTREL 660
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 601 LAPRSALALLEDTFDYPGRAHTFDDFCPCRCALGLQGCFAQFQSTVABLQRLKMKVGTREL 660
/ DB : : : : : : : : : : : : : : : : : : : : : : : : : :
/
/ RESULT 7
/ US-10-165-868-20
/ Sequence 20, Application US/10165868
/ Publication No. US20030143241A1
/ GENERAL INFORMATION:
/ APPLICANT: Reyes, Gregory R.
/ Twu, Jr-Shin
/ Purdy, Michael A.
/ Tam, Albert W.
/ Krawczynski, Krzysztof Z.
/ TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dehlinger & Associates
/ STREET: 350 Cambridge Avenue, Suite 250
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/165,868
/ FILING DATE: 06-Jun-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/870,985A
/ FILING DATE: 20-APRIL-1992
/ APPLICATION NUMBER: US 822,335
/ FILING DATE: 17-JAN-1992
/ APPLICATION NUMBER: US 505,888
/ FILING DATE: 05-APRIL-1990
/ APPLICATION NUMBER: US 420,921
/ FILING DATE: 13-OCTOBER-1989
/ APPLICATION NUMBER: US 367,486
/ FILING DATE: 16-JUNE-1989
/ APPLICATION NUMBER: US 336,672
/ FILING DATE: 11-APRIL-1989
/ APPLICATION NUMBER: US 208,997
/ FILING DATE: 17-JUNE-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 4600-0093.30
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 659 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/
```


Best Local Similarity 92.1%; Pred. No. 2e-256; Matches 607; Conservative 26; Mismatches 26; Indels 0; Gaps 0;	
QY	1 MRPRILLLLMLFPLPAPPPGQSGRRRRSGGGGFWGDRVDSQPPAIPYIHPTN 60
Db	1 MRPRALLLLFLVLLPMLPAPAGQPSGRRRRSGGTGGGFWGDRVDSQPPAIPYIHPTN 60
QY	61 PFAPDVTAAGAGPRVROPAPPLGSAWRDQORPAVASRRRPTTAGAAPLTAVAPADTP 120
Db	61 PFAADVSPQAGARPRQPPRPLGSAWRDQORPSAAPRRRSAPAGAAPLTAVSPADTA 120
QY	121 PVPDVSREGAILRROYNLSTSLTSSVATGNTLVLYAAPLSPLLPLOGDGTNTHIMATEAS 180
Db	121 PVPDVSREGAILRROYNLSTSLTSSVAGSNTLVLYAAPLPLLPLOGDGTNTHIMATEAS 180
QY	181 NYAQVRARATIRYRPLVNAVGGYAIISFWPQTITPTSDVMNSITSTDVRIIVQPGI 240
Db	181 NYAQVRVVRATIRYRPLVNAVGGYAIISFWPQTITPTSDVMNSITSTDVRIIVQPGI 240
QY	241 ASELVIPSERLHYRNOGWSRVETSGVAREEATSGLVMLCIHGSLVNSYNTPTTCALGLL 300
Db	241 ASELVIPSERLHYRNOGWSRVETTCVAREEATSGLVMLCIHGSPVNSYNTPTTCALGLL 300
QY	301 DPALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELTTTAAATRFMKDLYFTSTNG 360
Db	301 DPALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELTTTAAATRFMKDLHFTGTNG 360
QY	361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
Db	361 VGEVGRGIALTLFNLADTLGLLPTLTISSAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
QY	421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVL RANDVWLSTAAEY 480
Db	421 QDKGITIIPHDIDLGSRVVIQDYDNOHQDRPTSPAPSRPFSVL RANDVWLSTAAEY 480
QY	481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTL DGRPLSTIQYSKTFYVLP 540
Db	481 DQTYGSSTNPWVSDTVTLNVNATGAQAVARSLDWSKVTL DGRPLTTIQYSKTFYVLP 540
QY	541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db	541 LRKGLSFWEAGTTKAGYPYNYNTTASDQILIENAAAGHRVAISTYTTSLGAGPTSISAVGV 600
QY	601 LAPHSALEDDTDYPARAHTFDDFCPECRPLGQGCAPQSTVAELORLKMVKKTRE 659
Db	601 LAPHSALEDDTDYPARAHTFDDFCPECRTLGQGCAPQSTIAELORLKMVKKTRE 659
RESULT 10	
US-10-239-090A-53	
; Sequence 53, Application US/10239090A	
; Publication No. US20040101820A1	
; GENERAL INFORMATION:	
; APPLICANT: KABUSHIKI KAISHA TOSHIBA	
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co	
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,	
; TITLE OF INVENTION: detecting hepatitis E virus using the same	
; FILE REFERENCE: 0280741P	
; CURRENT APPLICATION NUMBER: US/10/239,090A	
; CURRENT FILING DATE: 2003-07-24	
; PRIOR APPLICATION NUMBER: JP2001-191837	
; PRIOR FILING DATE: 2001-06-25	
; NUMBER OF SEQ ID NOS: 57	
; SEQ ID NO 53	
; LENGTH: 660	
; TYPE: PRT	
; ORGANISM: Hepatitis E Virus JMY-Haw (ORF2)	
US-10-239-090A-53	
Query Match 93.1%; Score 3174; DB 16; Length 660;	
Best Local Similarity 92.1%; Pred. No. 2.9e-256;	
Matches 607; Conservative 25; Mismatches 27; Indels 0; Gaps 0;	

QY	1 MRPRILLLLMLFPLPAPPPGQSGRRRRSGGGGFWGDRVDSQPPAIPYIHPTN 60
Db	1 MRPRALLLLFLVLLPMLPAPAGQPSGRRRRSGGTGGGFWGDRVDSQPPAIPYIHPTN 60
QY	61 PFAPDVTAAGAGPRVROPAPPLGSAWRDQORPAVASRRRPTTAGAAPLTAVAPADTP 120
Db	61 PFAADVSPQAGARPRQPPRPLGSAWRDQORPSAAPRRRSAPAGAAPLTAVSPADTA 120
QY	121 PVPDVSREGAILRROYNLSTSLTSSVATGNTLVLYAAPLSPLLPLOGDGTNTHIMATEAS 180
Db	121 PVPDVSREGAILRROYNLSTSLTSSVAGSNTLVLYAAPLPLLPLOGDGTNTHIMATEAS 180
QY	181 NYAQVRARATIRYRPLVNAVGGYAIISFWPQTITPTSDVMNSITSTDVRIIVQPGI 240
Db	181 NYAQVRVVRATIRYRPLVNAVGGYAIISFWPQTITPTSDVMNSITSTDVRIIVQPGI 240
QY	241 ASELVIPSERLHYRNOGWSRVETSGVAREEATSGLVMLCIHGSLVNSYNTPTTCALGLL 300
Db	241 ASELVIPSERLHYRNOGWSRVETTCVAREEATSGLVMLCIHGSPVNSYNTPTTCALGLL 300
QY	301 DPALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELTTTAAATRFMKDLYFTSTNG 360
Db	301 DPALEFRNLTPGNTNTRVSRYSSTARHRLRRGADGTAELTTTAAATRFMKDLHFTGTNG 360
QY	361 VGEIGRGIALTLFNLADTLGLLPTLTISSAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
Db	361 VGEVGRGIALTLFNLADTLGLLPTLTISSAGGQLFYSRPVVSANGPTVKLYTSVENAQ 420
QY	421 QDKGIAIPHDIDLGESRVVIQDYDNOHQDRPTSPAPSRPFSVL RANDVWLSTAAEY 480
Db	421 QDKGITIIPHDIDLGSRVVIQDYDNOHQDRPTSPAPSRPFSVL RANDVWLSTAAEY 480
QY	481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTL DGRPLSTIQYSKTFYVLP 540
Db	481 DQTYGSSTNPWVSDTVTLNVNATGAQAVARSLDWSKVTL DGRPLTTIQYSKTFYVLP 540
QY	541 LRKGLSFWEAGTTKAGYPYNYNTTASDQLLVENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db	541 LRKGLSFWEAGTTKAGYPYNYNTTASDQILIENAAAGHRVAISTYTTSLGAGPTSISAVGV 600
QY	601 LAPHSALEDDTDYPARAHTFDDFCPECRPLGQGCAPQSTVAELORLKMVKKTRE 659
Db	601 LAPHSALEDDTDYPARAHTFDDFCPECRTLGQGCAPQSTIAELORLKMVKKTRE 659
RESULT 11	
US-10-239-090A-55	
; Sequence 55, Application US/10239090A	
; Publication No. US20040101820A1	
; GENERAL INFORMATION:	
; APPLICANT: KABUSHIKI KAISHA TOSHIBA	
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co	
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same, i	
; TITLE OF INVENTION: detecting hepatitis E virus using the same	
; FILE REFERENCE: 0280741P	
; CURRENT APPLICATION NUMBER: US/10/239,090A	
; CURRENT FILING DATE: 2003-07-24	
; PRIOR APPLICATION NUMBER: JP2001-191837	
; PRIOR FILING DATE: 2001-06-25	
; NUMBER OF SEQ ID NOS: 57	
; SEQ ID NO 55	
; LENGTH: 674	
; TYPE: PRT	
; ORGANISM: Hepatitis E Virus JKK-Sap (ORF2)	
US-10-239-090A-55	
Query Match 92.9%; Score 3166; DB 16; Length 674;	
Best Local Similarity 92.0%; Pred No. 1.4e-255;	
Matches 606; Conservative 24; Mismatches 29; Indels 0; Gaps 0;	
QY	1 MRPRILLLLMLFPLPAPPPGQSGRRRRSGGGGFWGDRVDSQPPAIPYIHPTN 60
Db	1 MRSRAFLFLVLLPMLPAPAGQPSGRRRRSGAGGFWGDRVDSQPPAIPYIHPTN 74

QY 61 PPAPDVTAAGAGRRVQPARPLGSARWDOAORPAVASRRRPTTAGAAPTAVAPADHTP 120
Db 75 PFASIDPAAAGAGRRVQPARPLGSARWDOAORPAVASRRRPTTAGAAPTAVAPADHTP 134
QY 121 PVPDVSRRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLPLLODGTNTHIMATEAS 180
Db 135 PVPDVSRRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLPLLODGTNTHIMATEAS 194
QY 181 NYAQYRVARATIRYRPLVNAVGGYAIISIFWFOPTTTTPTSDVMNSITSDVRILVQPGI 240
Db 195 NYAQYRVARATIRYRPLVNAVGGYAIISIFWFOPTTTTPTSDVMNSITSDVRILVQPGI 254
QY 241 ASELVIPSERLHYNQWRVSVEVSGVAEEATSGLVMLCTHGSVNSYTNTPYTGALGLL 300
Db 255 ASELVIPSERLHYNQWRVSVEVSGVAEEATSGLVMLCTHGSVNSYTNTPYTGALGLL 314
QY 301 DFALAEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 315 DFALAEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 374
QY 361 VGEIGRGIALTLFNADTLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 375 VGEVGRGIALTLFNADTLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 434
QY 421 QDKGIALPHDIDIGESRVVIQDYNQHEODRPTSPAPSPFPFVLRANDVWLSTAAEY 480
Db 435 QDKGIALPHDIDIGESRVVIQDYNQHEODRPTSPAPSPFPFVLRANDVWLSTAAEY 494
QY 481 DQSTYSSSTGPVVSVDVTLNVNATGAQAVARSLDWTKVLTDGRPLSTIOQYSKTFVLP 540
Db 495 DQTYSSSTNPVMYSDVTLVNATGAQAVARSLDWTKVLTDGRPLSTIOQYSKTFVLP 554
QY 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLIVENAAGHRVAISTYTTSLGAGPVSISAVV 600
Db 555 LRKGLSFWEAGTTKAGYPYNYNTASDQLIVENAAGHRVAISTYTTSLGAGPVSISAVV 614
QY 601 LAPHSALEEDTLDYPARAHTFDDFCPECRPLGQCAFQSTVAELQRLKMKVKGTRE 659
Db 615 LAPHSALEEDTLDYPARAHTFDDFCPECRPLGQCAFQSTVAELQRLKMKVKGTRE 673

RESULT 12

US-09-468-147-92
; Sequence 92, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/173,141
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 92
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 481
; OTHER INFORMATION: Xaa = Unknown or Other at position 542
; OTHER INFORMATION: Xaa = Unknown or Other at position 595

US-09-468-147-92

Query Match 92.7%; Score 3160; DB 10; Length 660;
Best Local Similarity 91.7%; Pred. No. 4.3e-255;
Matches 604; Conservative 27; Mismatches 28; Indels 0; Gaps 0;
QY 1 MRPRPILLLLMLPMLPAPPQPSGRRRGRSGSGGFWGDRVDSQPPAIPYIHPTN 60
Db 1 MRPRPILLLLMLPMLPAPPQPSGRRRGRSGSGGFWGDRVDSQPPAIPYIHPTN 60
QY 61 PFAPDVTAAGAGRRVQPARPLGSARWDOAORPAVASRRRPTTAGAAPTAVAPADHTP 120
Db 61 PFAPDVTAAGAGRRVQPARPLGSARWDOAORPAVASRRRPTTAGAAPTAVAPADHTP 120
QY 121 PVPDVSRRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLPLLODGTNTHIMATEAS 180
Db 121 PVPDVSRRGAILRRQYNLSPLTSSVATGTLNVLVYAAPLSPLPLLODGTNTHIMATEAS 180
QY 181 NYAQYRVARATIRYRPLVNAVGGYAIISIFWFOPTTTTPTSDVMNSITSDVRILVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGGYAIISIFWFOPTTTTPTSDVMNSITSDVRILVQPGI 240
QY 241 ASELVIPSERLHYNQWRVSVEVSGVAEEATSGLVMLCTHGSVNSYTNTPYTGALGLL 300
Db 241 ASELVIPSERLHYNQWRVSVEVSGVAEEATSGLVMLCTHGSVNSYTNTPYTGALGLL 300
QY 301 DFALAEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
Db 301 DFALAEFRNLTPGNTNTRVSRYSSTARHLRRGADGTAEITTTAATRFMKDLYFTSTNG 360
QY 361 VGEIGRGIALTLFNADTLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
Db 361 VGEVGRGIALTLFNADTLGGLPTELISAGGQLFYSRPVVSANGEPTVKLYTSVENAQ 420
QY 421 QDKGIALPHDIDIGESRVVIQDYNQHEODRPTSPAPSPFPFVLRANDVWLSTAAEY 480
Db 421 QDKGIALPHDIDIGESRVVIQDYNQHEODRPTSPAPSPFPFVLRANDVWLSTAAEY 480
QY 481 DQSTYSSSTGPVVSVDVTLNVNATGAQAVARSLDWTKVLTDGRPLSTIOQYSKTFVLP 540
Db 481 QDQTYSSSTNPVMYSDVTLVNATGAQAVARSLDWTKVLTDGRPLSTIOQYSKTFVLP 540
QY 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLIVENAAGHRVAISTYTTSLGAGPVSISAVV 600
Db 541 LRKGLSFWEAGTTKAGYPYNYNTASDQLIVENAAGHRVAISTYTTSLGAGPVSISAVV 600
QY 601 LAPHSALEEDTLDYPARAHTFDDFCPECRPLGQCAFQSTVAELQRLKMKVKGTRE 659
Db 601 LAPHSALEEDTLDYPARAHTFDDFCPECRPLGQCAFQSTVAELQRLKMKVKGTRE 659

RESULT 13
US-10-319-745-92
; Sequence 92, Application US/10319745
; Publication No. US20030211467A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/10/319,745
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: US/09/468,147A
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/173,141
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/061,199
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 92
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 481
; OTHER INFORMATION: Xaa = Unknown or Other at position 542
; OTHER INFORMATION: Xaa = Unknown or Other at position 595
US-10-319-745-92

Query Match          92.7%; Score 3160; DB 12; Length 660;
Best Local Similarity 91.7%; Pred. No. 4.3e-255;
Matches 604; Conservative 27; Mismatches 28; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLPMLPAPPQSGRRGRSGSGGGFWGDRVDSQFFAIPYIHPN 60
Db 1 MRPRVLLLLFLLPMLPAPPAGQSGRRGRSGGGFWGDRVDSQFFAIPYIHPN 60

QY 61 PFAPDVTAAGAGPRVRQPARPLGSAWRDQAPAVASRRRPTTGAAPLTAAPAHDP 120
Db 61 PFAADVSPQAGTRPRQPPRLGSAWRDQSKPSVAPRRSTPAGAAPLTAISPAPDTA 120

QY 121 PVPDVDSGAILRRQVNLSTSPLTSSVATGTLNLYAAPLSPLLPQDGTNTHIMATEAS 180
Db 121 PVPDVDSGAILRRQVNLSTSPLTSSVATGTLNLYAAPLSPLLPQDGTNTHIMATEAS 180

QY 181 NYAQYRVARATIRYRPLVNAVGVVAISISFWPQTTTTTSDVMSITSDVRIIVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGVVAISISFWPQTTTTTSDVMSITSDVRIIVQPGI 240

QY 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTTGALGLL 300
Db 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTTGALGLL 300

QY 301 DPALFERNLTPGNTNRVSRYSSTARHRLRGADGTAEITTAATRFMKDLYFTSTNG 360
Db 301 DPALFERNLTPGNTNRVSRYSSTARHRLRGADGTAEITTAATRFMKDLYFTSTNG 360

QY 361 VGEIGRGIALTLFNLADTLGLPTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLPTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420

QY 421 QDKGTAIPHDIIDGSRVVIQDYDNQHEQDRTPSPAPSRPFSVLRANDVWLSITAABY 480
Db 421 QDKGTAIPHDIIDGSRVVIQDYDNQHEQDRTPSPAPSRPFSVLRANDVWLSITAABY 480

QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIIQQYSKTFVFLP 540
Db 481 XQTTYGSSTNPMYVSDSVTLNVNATGAQAVARSLDWSKVTLGRPLSTIIQQYSKTFVFLP 540

QY 541 LRGLKLSFWEAGTTKAGYPYNNYNTASDQLLENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRGLKLSFWEAGTTKAGYPYNNYNTASDQLLENAAAGHRVAISTYTTSLGAGPVSISAVAV 600

QY 601 LAPHSALEDDTDYPARAHTFDDFCPCRLGLQGCAGFQSTVAELQRLKMKVKKTRE 659
Db 601 LAPHSALEDDTDYPARAHTFDDFCPCRLGLQGCAGFQSTVAELQRLKMKVKKTRE 659

RESULT 14
US-09-468-147-167
; Sequence 167, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauer, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushanwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
```

```
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/173,141
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 167
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 84
; OTHER INFORMATION: Xaa = Unknown or Other at position 230
; OTHER INFORMATION: Xaa = Unknown or Other at position 447
US-09-468-147-167
```

```
Query Match          92.5%; Score 3151; DB 10; Length 660;
Best Local Similarity 91.5%; Pred. No. 2.4e-254;
Matches 603; Conservative 26; Mismatches 30; Indels 0; Gaps 0;

QY 1 MRPRPILLLLMLPMLPAPPQSGRRGRSGSGGGFWGDRVDSQFFAIPYIHPN 60
Db 1 MRPRVLLLLFLLPMLPAPPAGQSGRRGRSGGGFWGDRVDSQFFAIPYIHPN 60

QY 61 PFAPDVTAAGAGPRVRQPARPLGSAWRDQAPAVASRRRPTTGAAPLTAAPAHDP 120
Db 61 PFAADVSPQAGTRPRQPPRLGSAWRDQSKPSVAPRRSTPAGAAPLTAISPAPDTA 120

QY 121 PVPDVDSGAILRRQVNLSTSPLTSSVATGTLNLYAAPLSPLLPQDGTNTHIMATEAS 180
Db 121 PVPDVDSGAILRRQVNLSTSPLTSSVATGTLNLYAAPLSPLLPQDGTNTHIMATEAS 180

QY 181 NYAQYRVARATIRYRPLVNAVGVVAISISFWPQTTTTTSDVMSITSDVRIIVQPGI 240
Db 181 NYAQYRVARATIRYRPLVNAVGVVAISISFWPQTTTTTSDVMSITSDVRIIVQPGI 240

QY 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTTGALGLL 300
Db 241 ASELVIPSERLHYRQWRSVETSGVAEEATSGLVMLCIHGSVNSYNTPTTGALGLL 300

QY 301 DPALFERNLTPGNTNRVSRYSSTARHRLRGADGTAEITTAATRFMKDLYFTSTNG 360
Db 301 DPALFERNLTPGNTNRVSRYSSTARHRLRGADGTAEITTAATRFMKDLYFTSTNG 360

QY 361 VGEIGRGIALTLFNLADTLGLPTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420
Db 361 VGEIGRGIALTLFNLADTLGLPTELISAGGQLFYSPVVSANGEPVTKLYTSVENAQ 420

QY 421 QDKGTAIPHDIIDGSRVVIQDYDNQHEQDRTPSPAPSRPFSVLRANDVWLSITAABY 480
Db 421 QDKGTAIPHDIIDGSRVVIQDYDNQHEQDRTPSPAPSRPFSVLRANDVWLSITAABY 480

QY 481 DQSTYGSSTGPVYVSDSVTLNVNATGAQAVARSLDWTKVTLGRPLSTIIQQYSKTFVFLP 540
Db 481 DQTTYGSSTNPMYVSDSVTLNVNATGAQAVARSLDWSKVTLGRPLSTIIQQYSKTFVFLP 540

QY 541 LRGLKLSFWEAGTTKAGYPYNNYNTASDQLLENAAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db 541 LRGLKLSFWEAGTTKAGYPYNNYNTASDQLLENAAAGHRVAISTYTTSLGAGPVSISAVAV 600

QY 601 LAPHSALEDDTDYPARAHTFDDFCPCRLGLQGCAGFQSTVAELQRLKMKVKKTRE 659
Db 601 LAPHSALEDDTDYPARAHTFDDFCPCRLGLQGCAGFQSTVAELQRLKMKVKKTRE 659

RESULT 15
US-10-319-745-167
; Sequence 167, Application US/10319745
; Publication No. US20030211467A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/10/319,745
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: US/09/468,147A
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/173,141
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/061,199
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 167
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 84
; OTHER INFORMATION: Xaa = Unknown or Other at position 230
; OTHER INFORMATION: Xaa = Unknown or Other at position 447
US-10-319-745-167

Query Match          92.5%; Score 3151; DB 12; Length 660;
Best Local Similarity 91.5%; Pred. No. 2.4e-254;
Matches 603; Conservative 26; Mismatches 30; Indels 0; Gaps 0;

Qy      1 MRPRPILLLLMLPAPPQPSGRRRGRSGGGGFWGDRVDSQPFAPYIHPN 60
Db      1 MRPRVLLLLFVLLPMLPAPPAGQPSGRRRGRSGGAGGFWGDRVDSQPFALPYIHPN 60

Qy      61 PEAPDVTAAAGAPRVQAPPLGSAWEDQAPAVASRRPTTAGAAPTAVAPAHDT 120
Db      61 PFAADVVSQAGAGPRPPRPPLXSAWSDQSRPSAAPRRRSAPAGAAPLTAVSPAPDTA 120

Qy      121 PVPDVSRGAILRRQYNLSPTSSVATGTLNLVLAAPLSPLLPLODGTNTHIMATEAS 180
Db      121 PVPDVSRGAILRRQYNLSPTSSVASGTNLVLAAPLNPLLPLODGTNTHIMATEAS 180

Qy      181 NYAQYRVARATIRYRPLVPNAVGYAISFWPQTTTTPTSDVMNSITSTDVRLVQPGI 240
Db      181 NYAQYRVVATIRYRPLVPNAVGYAISFWPQTTTTPTSDVMNSITSDVRLVQPGI 240

Qy      241 ASELVTPSERLHYRNOGWRSEVTSVGAEEATSGVLMCLTHGSLVNSYTNTPYTGALGLL 300
Db      241 ASELVTPSERLHYRNOGWRSEVITGVAAEEATSGVLMCLTHGSPVNSYTNTPYTGALGLL 300

Qy      301 DFALELEFRNLTPGNTNTRYSYSTARHLRERGADGTAEITTTAATRFMKDLYFTSTNG 360
Db      301 DFALELEFRNLTPGNTNTRYSYSTARHLRERGADGTAEITTTAATRFMKDLHFAGTNG 360

Qy      361 VGEIGRGIALTLNLAADTLGGLPTELISAGGQLFYSRPVSVSANGEPYVKLYTSVENAQ 420
Db      361 VGEVGRGIALTLNLAADTLGGLPTELISAGGQLFYSRPVSVSANGEPYVKLYTSVENAQ 420

Qy      421 QDKGIATPHDIDIGESRVVICDYNQHEQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480
Db      421 QDKGITIPHDIDIGDSRWIJDYDNOXEQDRPTSPAPSRPFSVLRANDVWLSTAAEY 480

Qy      481 DQSTYSGSTGPEVYVSDSVTLNVNATGAQAVARSLDWTKVTLDGRPLSTIQOYSKTFVLP 540
Db      481 DQTTYGSSITNPMYVSDVTLNVNATGAQAVARSLDWSKVTLDGRPLTTIQOYSKTFVLP 540

Qy      541 LRKGLSFEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLGAGPVSISAVAV 600
Db      541 LRKGLSFEAGTTKAGYPYNYNTTASDQILLIENAAAGHRVAISTYTTSLGAGPVSISAVGV 600
```

Qy 601 LAPHSALALLEDTLDYPARAHTDDFCPCRPGLGQCAFSQSTVAELQRLKMKVGKTR 659
Db 601 LAPHSALAVLEDTLDYPARAHTDDFCPCRTGLGQCAFSQSTIAELQRLKMKVGKTR 659

Search completed: August 16, 2004, 13:41:00
Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:35:43 ; Search time 89.8179 Seconds
(without alignments)
12476.164 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129
Perfect score: 660
Sequence: 1 ATGCGCCCTCGGCTATTTT.....GTAAACTCGGAGTTGTAG 1983

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2987630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xl
-Q=/Cgn2_1/USPTO_spool_p/US09851410/runat_16082004_125716_9617/app_query.fasta_1.4238
-DB=A_Geneseq_29Jan04 -QMT=fastan -SUFFIX=oli.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410@cgn_1_1_141@runat_16082004_125716_9617 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A_Geneseq_29Jan04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	660	100.0	660	2 AAR38785	Aar38785 HEV ORF2
2	660	100.0	660	2 AAR39306	Aar39306 Burma str
3	660	100.0	660	2 AAR96089	Aar96089 Hepatitis
4	660	100.0	660	2 AAW35826	Aaw35826 Hepatitis
5	660	100.0	660	2 AAW80197	Aaw80197 Protein e
6	660	100.0	660	2 AAW93389	Aaw93389 Human HEV
7	660	100.0	660	3 AAB24120	Aab24120 Hepatitis
8	660	100.0	660	4 AAB62523	Aab62523 HEV-Burma
9	660	100.0	660	5 AAO15699	Aao15699 Hepatitis
10	660	100.0	660	7 ADD24376	Add24376 Hepatitis

11	559	84.7	660	2 AAW76369	Aaw76369 Hepatitis
12	559	84.7	660	2 AAW71210	Aaw71210 Protein e
13	549	83.2	549	2 AAR96091	Aar96091 Hepatitis
14	541	82.0	660	4 AAE12980	Aae12980 Hepatitis
15	540	81.8	540	2 AAR96101	Aar96101 Hepatitis
16	525	79.5	525	2 AAR96103	Aar96103 Hepatitis
17	512	77.6	660	2 AAR70323	Aar70323 Hepatitis
18	448	67.9	549	2 AAW76367	Aaw76367 Hepatitis
19	446	67.6	660	5 ABJ04803	Abj04803 Hepatitis
20	408	61.8	660	2 AAR91814	Aar91814 Hepatitis
21	408	61.8	660	2 AAW93386	Aaw93386 Human HEV
22	408	61.8	660	5 ABB81670	Abb81670 Hepatitis
23	408	61.8	660	6 ADA50060	Ada50060 Hepatitis
24	408	61.8	660	7 ADB97807	Adb97807 HEV ORF2
25	396	60.0	660	2 AAW93388	Aaw93388 Human HEV
26	327	49.5	327	2 AAR96093	Aar96093 Hepatitis
27	327	49.5	327	2 AAW35820	Aaw35820 Hepatitis
28	321	48.6	660	2 AAR51265	Aar51265 HEV strai
29	321	48.6	660	2 AAW81520	Aaw81520 Hepatitis
30	310	47.0	660	5 ABJ04816	Abj04816 Hepatitis
31	304	46.1	660	5 ABJ04820	Abj04820 Hepatitis
32	298	45.2	660	5 ABJ04811	Abj04811 Hepatitis
33	284	43.0	660	2 AAW93395	Aaw93395 Human HEV
34	264	40.0	660	2 AAW93392	Aaw93392 Human HEV
35	259	39.2	660	2 AAW93390	Aaw93390 Human HEV
36	255	38.6	660	5 ABJ04807	Abj04807 Hepatitis
37	251	38.0	660	5 ABJ04818	Abj04818 Hepatitis
38	251	38.0	660	5 ABJ04809	Abj04809 Hepatitis
39	250	37.9	660	5 ABJ04819	Abj04819 Protein e
40	228	34.5	660	2 AAR14619	Aar14619 Protein e
41	217	32.9	660	5 ABJ04822	Abj04822 Hepatitis
42	214	32.4	660	5 ABJ04821	Abj04821 Hepatitis
43	210	31.8	213	4 AAU00496	Aau00496 Hepatitis
44	210	31.8	214	5 ABB81671	Abb81671 Hepatitis
45	210	31.8	232	5 ABB81679	Abb81679 Hepatitis

ALIGNMENTS

RESULT 1
AAR38785
ID AAR38785 standard; protein; 660 AA.
XX
AC AAR38785;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 11-JAN-1994 (first entry)
XX
DE HEV ORF2 protein.
XX
KW Enterically transmitted non-A non-B hepatitis; ET-NANB; vaccine.
XX
OS Hepatitis E virus; Burma strain.
XX
FH Key Location/Qualifiers
FT Peptide 225..660
FT Peptide /label= C2 334..660
FT Peptide /label= SG3 613..660
FT Peptide /label= 406.3-2
XX
PN WO9314116-A1.
XX
PD 22-JUL-1993.
XX
PF 15-JAN-1993; 93WO-US000459.
XX
PR 17-JAN-1992; 92US-00822335.
PR 01-MAY-1992; 92US-00876941.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.

```
PA (USSH ) US SEC DEPT HEALTH.
XX
PI Reyes GR, Bradley DW, Tam AW, Carl M;
XX
XX WPI; 1993-243144/30.
XX DR N-PSDB; AAQ47129.
XX
XX New immunogenic hepatitis E virus (HEV) peptides (s) - are from the ORF1,
PT ORF2 and ORF3 regions of HEV, useful as a vaccine against HEV infection.
XX
XX Disclosure; Fig 8; 48pp; English.
XX
XX Immunogenic hepatitis E virus (HEV) peptides are selected from the ORF1,
CC ORF2 and ORF3 regions of HEV. The peptides can be used in vaccines to
CC prevent infection by HEV. The antibodies can neutralise and block HEV
CC infection and can be used to prevent or treat HEV infection. The peptides
CC and antibodies can also be used as diagnostic reagents. (Updated on 25-
CC MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX
XX Sequence 660 AA;
SQ
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAR38785 (1-660)
QY 1 ATCGCGCTCGGCTATTTGTGCTCTCTCATGTTTGGCTATGTCGCGCGCCA 60
Db 1 MetArgProArgProLeuLeuLeuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCGCGGTGACGGGTGATTTCTAGCCCTTCGCAATCCCTTATATTCATCCAAAC 120
Db 21 ProProGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlySerGlyGly 40
QY 121 TTCTGGGTGACGGGTGATTTCTAGCCCTTCGCAATCCCTTATATTCATCCAAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCCGATGTCACGGTGGCGCGGGCTGGACCTCGTGTTCGCCAACCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyValArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCCGCTTGGCGGTGACGAGCCCGAGCGCCCGCGTTCGCTCACGTCGT 300
Db 81 ArgProLeuGlySerAlaTTPArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACCTTACACAGTGGGGCGGCGCTAACCGCGGTGCTCGGCGCCATGACACCCCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTCGACTCCCGCGGCGCATCTTGCGCGGAGTATACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCGGCGCACCGGCACTAACCTGGTCTCTTATGCGCGCCCTCT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTACCTTTAGGACGGCACCAATACCCATATATATGCGCACGGAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTTATGCCAGTACCGGGTTCGCCGTGCCAATCCGTTTACGCCCGCTGGTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTGCGCGGTACGCCATCTCCATCTCATTTCTGGCCACAGACCAACCCCGGACG 660
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
661 TCGGTGATATGAATTCATTAACCTCGACGGATGTTGCTATTTAGTCCAGCCGCGCAT 720
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
721 GCTCTGAGCTTGATCCCAAGTGAAGCGCTACACTATCGTAACCAAGSCTCGCGTCC 780
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
781 GTCCAGACCTCTCGGGTGGCTGAGGAGAGGCTACCTCTGGTCTTGTATGCTTTGCATA 840
261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
841 CATGGCTCACTGTAATTCCTATATACTATACACCTATACCGTCCGCTCGGGCTGTTG 900
281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeu 300
901 GACTTTGCCCTTGAGCTTGAGTTTCGCACTTACCCCGGTAAACCAATACCGCGGTC 960
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
961 TCCCGTTATTCAGCAGCTGCTGCCACCGCTTCGTCGCGGTGCGGAGCGGAGTCCGAG 1020
321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
1021 CTCACCAACACGCTGCTACCGCTTATGAAGACCTCTATTTACTAGTACTAATGGT 1080
341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTGTGTGACACTCTGCTT 1140
361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
1141 GCGCGCTCCGACAGAAATGATTCGTGGCTGGTGGCGAGTGTCTACTCCGTCCTC 1200
381 GlyGlyLeuProThrGluLeuLeuSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
1201 GTTGTCTCAGCAATCGCGAGCGCTGTTAAGTGTATATCTATATCTAGAGATGCTCAG 1260
401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
1261 CAGGATAAGGATTCGAATCCCGCATGACATTCGAGCTCGGAGAACTCTCGTGGTTATT 1320
421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
1321 CAGGATTTATGAATACCAACATGAACAGATCGCGCGAGCGCTTCTCCAGCCCATCGCGC 1380
441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
1381 CCTTCTCTGCTCTCGAGCTAATGATGTGCTTGGCTCTCTCTCCAGCTGCGGAGTAT 1440
461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGlyTyr 480
1441 GACCACTCACTTATGCTCTTCGACTGGCCCGCTTATGTTTCTGACTCTGTGACCTTG 1500
481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
1501 GTTAATGTCGAGACCGCGCGGAGCGGCTGCGCGGTGCGCTCGATTGGACCAAGTCCACA 1560
501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
1561 CTTGAGCGTCCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTCTGCGC 1620
521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
1621 CTCGCGGGTAAAGCTCTCTTTCTCGGAGGCGAGGACCAACTAAAGCCGGGTACCTTATAAT 1680
541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
1681 TATAACCACTCTCTAGCGACCACTGCTTGTGAGAAATCGCGCGGACCGGGTCTGCT 1740
```

Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCACTAGCTGGTGGTCTCCGCTCCATTTCTGCGGTTCCCGTT 1800
Db 581 IleSerThrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGACTACCTCCCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGATGATTTCTGCGCAGAGTCCCGCCCTTGCGCTTTCAGGGCTGCGCTTC 1920
Db 621 HisThrPheAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGCTACTCTGCTGAGCTTCAGCGCTTACAGTGAAGTGGTAAACTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuGlyMetLysValGlyLysThrArgGluLeu 660
RESULT 2
AAR39306
ID AAR39306 standard; protein; 660 AA.
XX
AC AAR39306;
XX
DT 25-MAR-2003 (revised)
DT 14-FEB-1994 (first entry)
XX
DE Burma strain HEV ORF2 putative virus capsid protein.
XX
KW Hepatitis E virus; vaccine; neutralising antibodies; infection; block;
KW open reading frame; antibodies.
XX
OS Hepatitis E virus.
XX
XX WO9314208-A2.
XX
XX 22-JUL-1993.
XX
XX 19-JAN-1993; 93WO-US000475.
XX
XX 17-JAN-1992; 92US-00822335.
PR 20-APR-1992; 92US-00870985.
XX
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Reyes GR, Bradley DW, Twu J, Purdy MA, Tam AW, Krawczynski KZ;
XX
XX WPI; 1993-243223/30.
DR N-PSDB; AAQ46813.
XX
XX Antigen and antibody vaccines against hepatitis E virus infection -
PT contain peptide(s) derived from capsid protein C-terminal or antibodies
PT against protein.
XX
XX Disclosure; Fig 7; 43pp; English.
XX
XX The sequence is that of the putative virus capsid protein encoded by
CC Burma strain hepatitis E virus (HEV) open reading frame ORF2. This
CC protein or peptide fragments of it may be used in a vaccine composition
CC for immunising an individual against HEV. Antibodies raised against these
CC peptides can also be used in such vaccines. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
XX Sequence 660 AA;
SQ

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-851-410a-6_COPY_5147_7129 (1-1983) x AAR39306 (1-660)

QY 1 ATGCGCCCTCGCCCTATTTTGTGTGCTCTCATGTTTTCCTATGCTGCCCGCGCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCCCGGCTCAGCGCTCTGCGCGCGTCTGTTGGCGCGCGCAGCGGCTTCCCGCGGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40
QY 121 TTTCTGGGCTGACCGGGTGAATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCTTTCGCCCGCGATGTCAACCGCTGCGCGCGGGCTGAGACTCTGTGTTGCGCAACCCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCCGCTTGGCTGACCGAGCGCCGCGCGCGCGCTTGCCTCAGCTCGT 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTACACAGCTGGCGCGCGCTAACCGCGGTCTCGCTCCGCGCCCATGACACCCCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CGAGTCTGTATGCTGACTCCCGCGCGCGCTTCTTGGCGCGCGAGTATAACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACTCTCTCGTGGCGCACCGGCACTAACCTGCTTCTTATGCGCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCTTTTACCCCTTTCAGGACCGCACCAATACCCATATAATGCCACGGAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTCCCGTCCCAATCCCTTACCGCCCGCTGCTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGGGTTAGCCCTCTCCATCTCATCTGCGCACAGACACACACCCCGGAG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATTAACCTCGAGCGATGTTCTGTTATTTAGTCCAGCCCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTGAAGCGCTTACACTATCGTAAACCAAGGTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACTCTGGGGTGGCTGAGAGGAGGCTACTCTGCTGTTTATGCTTGTATGCTTGCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCAGCTCGTAAATTCCTATATAACACCTATACCGGTGCGCTCGGGGTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTCCTTGTGAGCTTGTGCAACCTTACCCCGCGTAAACACCAATACCGGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGTTATTCAGCACTGCTCGCCACCGCTTGTGCGGTGGGAGGAGCTGCCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGGCTGCTACCGGCTTTTATGAAGGACCTCTATTTTACTAGTACTAATCGT 1080

QY 361 CCAGTGCCTGATGTCGACTCCCGCGCGCCACTTTCGCGCGCAGTATACCTATCAACA 420
DB 121 ProValProAlaSerValAspSerArgGlyAlaLeuAraGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACTCTTCCTGCGCCACCGCACTAACCTGGTTCCTTATGCGCGCCCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGGCTTTACCTCCCTCAGACCGCACCATATCCATATATATGCGCACCGAAGCTTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATATGCCAGTACCGGCTGCCCGTCCCAATCCGTTACCGCCGCTGGTCCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaAlaThrIleAraGlyTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGCGGTACGCCATCTCCATCTCATCTCTGGCCACAGACACCAACCCGAGC 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATCAATCACTCGAGCGATGTCGTATTTAGTCAGCGCGGCATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValAraGlyLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTCAGCGCTACACTATCGTAACCAAGGCTGGCGCTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
QY 781 CTCGAGACCTCTGGGCTGGTGAGGAGGCTACCTCTGTCCTTGTATGCTTTGCATA 840
DB 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCACTCGTAATCTTACTACTATACACCTATACCGTTCACCGCTCGGCTGTG 900
DB 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGACTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCAGCACTGCTCGCACCGCTTCGTCGCGGTGCGGAGCGGACTCCGAG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisArgLeuAraGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACGCGCTACCGCTTATGAGGACCTCTATTTACTAGTACTATGCT 1080
DB 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 CTCGCTGAGATCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGCTGACACTCTGCTT 1140
DB 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCGACAGAAATGATTTGCTGCGGTGTCGCGCTGCTGCTTACTCCCGTCC 1200
DB 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGAGTTCGTCGCGCTGTCGCGCTGTCGCTGCTGCTGCTGCTGCT 1260
DB 401 ValValSerAlaAsnGlyGlnProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAGGGTATTCGAATCCCGATGACATTCGAGTTCGAGAGTCTCTGCTGGTATT 1320
DB 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGlnSerArgValIle 440
QY 1321 CAGGATATGATACCAACATCAAGATCGCGCAGCGCTTCTCCAGCCCATCGCGC 1380
DB 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCGCGATAT 1440
DB 461 ProPheSerValLeuAraAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCAGTCCACTTATGGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500

DB 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAAATGTTGACACCGCGCGAGCGCTTCCCGCTCGCTCGATTCGACCAAGGTCA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTGACGGTCCGCCCTCTCCACCATCCAGAGTACTCGAAGACCTCTCTTCTGCTCGCG 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAAAGCTCTCTTCTGGGAGGACGACCACTAAAGCGCGGTACCTTATAAT 1680
DB 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACACTGCTAGGACCACTGCTTCGAGATGCGCGCGGACCGGTGCT 1740
DB 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCACCTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB 581 IleSerThrThrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGAGTACCTGCGCGCGC 1860
DB 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCTTCAGGCTGCGCTTTC 1920
DB 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGCTGAGCTTACGCGCTTAAGTGAAGGTGGTAAACTCGGAGTTG 1980
DB 641 GlnSerThrValAlaGluLeuGlnArgLeuLysValGlyLysThrArgGluLeu 660

RESULT 4
AAW35826
ID AAW35826 standard; protein; 660 AA.
XX AAW35826;
AC AAW35826;
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 26-FEB-1998 (first entry)
DE Hepatitis E virus Burma strain protein from ORF2.
XX Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody;
KW diagnosis; HEV.
XX Hepatitis E virus; - Burma strain.
XX US5686239-A.
PN 11-NOV-1997.
XX 09-MAY-1994; 94US-00240049.
XX 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-00336672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 05-APR-1991; 91US-00681078.
PR 17-JAN-1992; 92US-00822335.
PR 20-APR-1992; 92US-00870985.
PR 01-MAY-1992; 92US-00876941.
XX (GENE-) GENELABS TECHNOLOGIES INC.
PA Tam AW, Reyes GR, Yarbough PO;
XX WPI; 1997-558132/51.
DB

DR N-PSDB; AAT96959.
XX Diagnosis of hepatitis E virus Burma and Mexico strain infection - by
PT immunoassay with hepatitis E virus peptide antigens.
XX
XX Disclosure; Col 39-42; 36pp; English.

XX A method has been developed for detecting hepatitis E virus (HEV)
CC antibodies (Ab). The method comprises: (a) reacting a serum sample with a
CC HEV peptide antigen; and (b) examining the peptide for the presence of
CC bound Ab, where the presence of bound Ab indicates the presence of HEV
CC Ab. The present sequence represents the protein from the open reading
CC frame, ORF2, from HEV Burma strain. The method can be used to diagnose
CC infection with the enterically transmitted non-A/non-B viral hepatitis
CC agent HEV, specifically the HEV Burma and Mexico strains. (Updated on 25-
CC MAR-2003 to correct PF field.) (Updated on 17-OCT-2003 to standardise OS
CC field)

XX Sequence 660 AA;

Alignment Scores:

Pred. No.:	0	Length:	660
Score:	660.00	Matches:	660
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAW35826 (1-660)

QY	1	ATCGCGCTCGGCTATTTGTCCTCTCATGTTTGGCTATGTCGCCGCGCCA	60
Db	1	MetArgProArgProIleLeuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20
QY	61	CCGCGCGCTCAGCGCTGCGCGCTGCGCGCGCGCGCGCGCGGTCCTCGCGGGT	120
Db	21	ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly	40
QY	121	TTCTGGGTCAACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC	180
Db	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn	60
QY	181	CCCTTCGCCCGGATGTCACCGCTGCGCGCGCGGCTGGACCTCGTGTGCCAACCGCC	240
Db	61	ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla	80
QY	241	CGACCACTCGGCTCCGCTGGCGTGCACCGCCAGCGCCCGCGCTGCTCAGCTCGT	300
Db	81	ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg	100
QY	301	AGACCTTACCACAGCTGGGCGCGCGCTAACCGCGCTCGCTCCGCGCCATGACACCCCG	360
Db	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro	120
QY	361	CCAGTCCGCTGATGTCAGCTCCCGCGCGCATCTTCCGCGCGGAGTATACCTATCAACA	420
Db	121	ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr	140
QY	421	TCTCCCTTACCTCTTCGTTGGCCACCGCACCTAACTGCTCTTTATGCGCGCTCTT	480
Db	141	SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu	160
QY	481	AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATATATGCGCACGGAAGCTTCT	540
Db	161	SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180
QY	541	AATTATGCCAGTACCGGTTGCGGTGCCACAAATCGGTTACCGCGCTGTCGCCAAT	600
Db	181	AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn	200
QY	601	GCTGTGCGGGTTACGSCATCTCATCTATTCTGGCCACAGACCAACCCCGGACG	660
Db	201	AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr	220

QY	661	TCGTTGATATGAATTCATATACTCGACGGATGTTCTGTTATTTTAGTCCAGCCGCGCAT	720
Db	221	SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle	240
QY	721	GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATCGTAAACCAAGCTGCGCTCC	780
Db	241	AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer	260
QY	781	GTCGAGACCTCTGGGTGCTGAGGAGGAGGCTACCTCTGCTGCTTGTATGCTTGCAT	840
Db	261	ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280
QY	841	CATGGCTCACTCGTAAATTCCTATCTAATACACCTATACCGGTGCGCTCGGCTGTTG	900
Db	281	HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu	300
QY	901	GACTTTGCCCTTGAGCTTGAGTTTCGAACCTTTACCCCGGTAAACCAATACGCGGTC	960
Db	301	AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal	320
QY	961	TCCCGTTATTCAGCACTGCTCGCACCGCTTCGTCGGGTGGGACGGAGCTGCCGAG	1020
Db	321	SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu	340
QY	1021	CTCACCAACGCGCTGCTACCCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGT	1080
Db	341	LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly	360
QY	1081	GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTTCACTTCTGCTGACCTCTGCTT	1140
Db	361	ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu	380
QY	1141	GCGGCGCTCGCAGACAGATTGTTTCGTCGGGTGGGCGGAGCTGTCTACTCCGCTCC	1200
Db	381	GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro	400
QY	1201	GTTGTCTCAGCAATGCGGAGCGCACTGTTTAACTTGTATACATCTGTAGAGAATGCTCAG	1260
Db	401	ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln	420
QY	1261	CAGGATAAGGTATTCGATCCCGCATGACATTGACCTCGGAGAACTCTGTGTGGTTATT	1320
Db	421	GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle	440
QY	1321	CAGGATTATGATAACCAACATGAAGATCGCGCAGCGCTTCTCAGCCCATCGCGC	1380
Db	441	GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg	460
QY	1381	CCTTTCTCTCTCGAGCTAATGATGCTTGGCTCTCTCTCACCGCTGCGGAGTAT	1440
Db	461	ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr	480
QY	1441	GACCACTCACTATGCTCTTCGACTGGCCCGAGTTATGTTTCTGACTCTGTGACCTTG	1500
Db	481	AspGlnSerThrTyrGlySerThrGlyProValTyrValSerAspSerValThrLeu	500
QY	1501	GTTAATGTTGCGACCGCGCAGCGGCTTGGCCGCTCGCTCGATTTGGACCAAGGTACA	1560
Db	501	ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspIleThrLysValThr	520
QY	1561	CTTGACGGTGGCCCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGCTCGCG	1620
Db	521	LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro	540
QY	1621	CTCCGCGGTAAAGCTCTCTTCTGGGAGCGAGGACCACTAAAGCCGGTACCTTTAAT	1680
Db	541	LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn	560
QY	1681	TATAACCACTGCTAGCGACCAACTGTTTGTGAGAATGCCCGCGGACCCGGGTGCT	1740
Db	561	TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla	580

QY 1741 ATTCCACTTACACACTAGCTGGGTGCTGGTCCCGTCTCCATTTCTTCGCGTTGCCGTT 1800
Db |||||||
581 IleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATGCTTGAGGATACCTTGAGTACCTGCCCGGCC 1860
Db |||||||
601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTGATGATTCTGCGCCAGAGTGCAGCCCTTGGCTTCAGGGCTGCGCTTTC 1920
Db |||||||
621 HisThrPheAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTCAGCTTTCAGGCGCTTAAGATGAAGGTGGTAAACTCGGAGTTG 1980
Db |||||||
641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 5
AAW80197
ID AAW80197 standard; protein; 660 AA.
XX AC AAW80197;
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 23-DEC-1998 (first entry)
XX DE Protein encoded by ORF2 of ET-NANB (HEV) Burma strain DNA sequence.
XX KW Enterically transmitted non A non B hepatitis virus; ET-NANB;
KW Hepatitis E virus; HEV; Burma HEV isolate; vaccine; diagnostic probe.
XX OS Non-A.
OS non-B hepatitis virus.
XX US5824649-A.
XX 20-OCT-1998.
XX 07-JUN-1995; 95US-00475807.
XX 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-00336672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 25-JUL-1994; 94US-00279823.
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX Bradley DW, Reyes GR, Krawczynski KZ, Tam A, Fry KE, Yarbough PO;
XX WPI; 1998-582599/49.
DR N-PSDB; AAV66321.
XX Hepatitis E virus proteins - useful for diagnosis or vaccine production
PT the virus.
XX Claim 6; Col 65-70; 47pp; English.
XX AAW80196-98 are encoded by the genome of the Burma strain of enterically
CC transmitted non A non B hepatitis virus (ET-NANB) (hepatitis E virus
CC (HEV)). The specification describes an isolated protein which is
CC specifically immunoreactive with antibodies present in individuals
CC infected with HEV and encoded by a sequence contained in an open reading
CC frame (ORF) of an HEV genome. The genome has a sequence that is more than
CC 70% identical to the ORF1 sequence from Burma HEV isolate. The protein is
CC used as a vaccine and a diagnostic probe for ET-NANB. (Updated on 25-MAR-
CC 2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 660 AA;
Alignment Scores:
Pred. No.: 0 Length: 660

Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAW80197 (1-660)
QY 1 ATGCGCCCTCGGCTCATTTTGTGTGCTCTCTCATGTTTTCCTATGCTGCGCGGCCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCCCCGCTCAGCCCTCTGCGCGCGCTGCGCGCGCGCAGCGCGGTTCGCCGCGTGT 120
Db |||||||
21 ProProGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlySerGlyGly 40
QY 121 TTCGGGTGACCCGGTGTATTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db |||||||
41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCCGATGTCACCGCTGCGCGCGCGGCTGGACCTCGTGTTCGCAACCCGCC 240
Db |||||||
61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTTGGCGTGACCAAGCGCCAGCGCGCGCTTCACGTGCT 300
Db |||||||
81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTACACAGCTGGCGCGCGCTTAACCGCGTGTCTCGCGCCCATGACACCCCG 360
Db |||||||
101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTCGACTCCCGCGCGCGCCATCTTGGCGCGCGCAGTATTAACCTACA 420
Db |||||||
121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCTGGCCACCGCGCACTAACCTGGTTCCTTATGCCGCCCTCTT 480
Db |||||||
141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCCCTTCAGSACGGCACCAATACCCATATAATGCCACGGAAGTTCT 540
Db |||||||
161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGCGGTGCGCGTCCCAATCCGTTACCGCGCGCTGCTCCCAAT 600
Db |||||||
181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGCGGTACGCCATCTCCATCTCATCTCTGGCCACAGACCACACCCCGCAGC 660
Db |||||||
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATGATCAATTCAATTAACCTCGACGAGTGTTCGTATTTAGTCCAGCCCGCAT 720
Db |||||||
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTACGCGCTACACTATCTGAACAGGCTGGCGCTCC 780
Db |||||||
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACTCTCGGGTGGCTGAGGAGGAGGTACTCTGGTCTGTGTTATGCTTTGCAT 840
Db |||||||
261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCACTCGTAAATTCCTATATAACCTATACCGGTGCGCTCGGGCTGTG 900
Db |||||||
281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGCGGTAAACCAATACGCGGTC 960
Db |||||||
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATCCAGCACTGCTCGCCACCGCTTCGTGCGGTGGGACGGGACTGCCGAG 1020

Db	81	ArgProLeuGlySerAlaIrrArgAspGlnAlaGlnArgProAlaValAlaSerArgArg	100	
QY	301	AGACCTACACAGCTGGGGCGCGCGCTAACCGGGTGGCTCCGGCCCATGACACCCCG	360	
Db	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro	120	
QY	361	CCAGTGCCTGATGTGCAGCTCCCGCGCGGCATCTTGGCGCGGCAGTATAACTATCAACA	420	
Db	121	ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr	140	
QY	421	TCTCCCTTACCTCTCCGTGGCCACCGCACATAACCTGGTCTTTATGCCGCCCTCTT	480	
Db	141	SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu	160	
QY	481	AGTCGGCTTTTACCCCTTCAGACCGCACCATATCCATATATATGAGCCACGGAAGCTTCT	540	
Db	161	SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180	
QY	541	AATTATGCCAGTACCGGGTTCCCGGTGCCACAATCCGTTACCGCGCTGGTCCCAAT	600	
Db	181	AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn	200	
QY	601	GCTGTCGGCGGTTACGCCATCTCCATCTCAATCTCGGCCACAGACACCAACCCCGCAG	660	
Db	201	AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr	220	
QY	661	TCCGTTGATATCAATTCATACCTCGAGGATGTTCGTATTTAGTCAGCCCGGCATA	720	
Db	221	SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle	240	
QY	721	GCCTCTGAGCTGTGATCCCAAGTCAGCGCCPACACTATCGTAACCAAGCTGGCGCTCC	780	
Db	241	AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer	260	
QY	781	GTCGAGACTCTGGGGTGGCTGAGAGAGGCTACTCTGCTGTGTATGCTTTGTCATA	840	
Db	261	ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280	
QY	841	CATGCTCACTCGTAATTCCTATACTATAACCTATACCGGTGCCCTCGGGCTGTG	900	
Db	281	HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu	300	
QY	901	GACTTTGGCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACGGGGTC	960	
Db	301	AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal	320	
QY	961	TCCCGTTATTCAGACTGCTCGCCACCGCCTTCGTCGGGTGGCGAGCGGACTCCGAG	1020	
Db	321	SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu	340	
QY	1021	CTCACACACCGGCTGCTACCGCTTATGAGGACCTCTATTTACTAGTACTAATGTT	1080	
Db	341	LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly	360	
QY	1081	GTCGCTGAGATCGCGCGGGTAGCCCTCACCCCTGTTCAACCTTGCTGACACTCTGCTT	1140	
Db	361	ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu	380	
QY	1141	GGCGGCTTCGCCACAGAAATTTGTCGGCTGGTGGCCAGCTGTTCTACTCCCGTCCC	1200	
Db	381	GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro	400	
QY	1201	GTTGCTCAGCAATAGCGGACCGGACTGTAGTTGTATACATCTGTAGAGATGCTCAG	1260	
Db	401	ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln	420	
QY	1261	CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAATCTCGTGTGGTTATT	1320	
Db	421	GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle	440	
QY	1321	CAGGATTATGATACCAACATGAACATCGCGGACGCGCTTCTCCAGCCCATCGCGC	1380	
Db	441	GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg	460	
QY	1381	CCTTTCTCTGCTTCGAGCTAAATGATGCTTTGGCTCTCTCTCACCGCTCCGAGTAT	1440	
Db	461	ProPheSerValLeuArgAlaAsnAspValLeuTyrPheLeuSerLeuThrAlaAlaGluTyr	480	
QY	1441	GACCAGTCCACTTATGGCTCTTCGACTGGCCAGCTTTATGTTTCTGACTCTCTGTGACCTTG	1500	
Db	481	AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu	500	
QY	1501	GTTAATGTTGGACACCGCGCGCAGCGCTGGTCCCGTGGCTCGATTGACCAAGGTACA	1560	
Db	501	ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLysValThr	520	
QY	1561	CTTGACGGTCCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTGCTCGCG	1620	
Db	521	LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro	540	
QY	1621	CTCCCGGCTAAGCTCTCTTTCTGGGAGGCAGGCACAACTAAAGCCGGGTACCTTATAT	1680	
Db	541	LeuArgGlyLysLeuSerPheTyrPheGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn	560	
QY	1681	TATAACACACTGTAGCGACCAACTGCTGTGAGAAATGCCCGCGGCACCGGTGCT	1740	
Db	561	TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla	580	
QY	1741	ATTTCACCTTACACCACTAGCTGGTGTGCTGCTCCGCTCCATTTCTCGGGTCCGCT	1800	
Db	581	IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal	600	
QY	1801	TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGGACTACCTCCCGCGCC	1860	
Db	601	LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla	620	
QY	1861	CATACCTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCCTTCAGGGCTGCGCTTTC	1920	
Db	621	HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe	640	
QY	1921	CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAGAGTGAAGTGGTAAACCTCGGAGTTG	1980	
Db	641	GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu	660	
RESULT 7				
ID	AAB24120	standard; protein; 660 AA.		
XX	AAB24120;			
AC	AAB24120;			
DT	29-JAN-2001	(first entry)		
XX	Hepatitis E virus Burma strain ORE2 protein sequence SEQ ID NO:8.			
DE	Hepatitis E virus; HEV; non-A/non-B viral hepatitis; viral protein;			
XX	immunological; diagnosis; hepatitis; infection; identification;			
KW	detection; immunoreactive; hepatotropic; antiinflammatory; virucide;			
KW	vaccine; antiviral; antigenic; antibody; antigen.			
XX	Hepatitis E virus.			
OS	US6120988-A.			
PN	19-SEP-2000.			
XX				
PD	07-JUN-1995;	95US-00478507.		
XX				
PF	17-JUN-1988;	88US-00208997.		
XX				
PR	11-APR-1989;	89US-00336672.		
XX				
PR	16-JUN-1989;	89US-00367486.		
XX				
PR	13-OCT-1989;	89US-00420921.		
XX				
PR	05-APR-1990;	90US-00505888.		
XX				
PR	05-APR-1991;	91US-00681078.		
XX				
PR	25-JUL-1994;	94US-00279823.		

PA (GENE-) GENELABS TECHNOLOGIES INC.
PA (USGO) US GOVERNMENT.

XX Yarbough PO, Krawczynski KZ, Fry KE, Bradley DN, Tam A, Reyes GS;
PI
DR WPI; 2000-593712/56.
DR N-PSDB; AAA99259.

Identifying recombinant antigen immunoreactive with antibody induced by hepatitis E virus (HEV), for detecting HEV infection, comprises immunoreacting a polypeptide from an HEV genome with an HEV-positive antiserum.

PS Claim 6; Col 63-68; 46pp; English.

The present invention describes a method for identifying a recombinant antigen immunoreactive with a hepatitis E virus (HEV)-induced antibody. The method comprises producing a polypeptide derived from an HEV genome, and immunoreacting the polypeptide with an HEV-positive antiserum and selecting the polypeptide as a recombinant antigen if the polypeptide reacts with the HEV-positive antiserum. The method is useful for identifying recombinant antigen immunoreactive with antibody induced by HEV. The enterically transmitted non-A/non-B hepatitis (ET-NANB (also known as HEV))-specific fragments are useful for identifying ET-NANB-derived cDNAs, which contain additional sequence information, as primers for detecting ET-NANB viral genomic material in a patient sample, for the synthesis of polypeptides for use in immunoassays, and for identifying similar antigenic regions encoded by related viral strains, e.g. Burmese strain. The antigens are especially useful in the preparation of vaccine against ET-NANB infection. These antigens may further be used to prepare antibodies to ET-NANB virus particles for use directly as antiviral agents, and to produce antiserum designed for pre- or post-exposure prophylaxis. The present sequence represents a specifically claimed HEV Burma strain protein sequence for use in the present invention.

Sequence 660 AA;

Alignment Scores:

	0	660.00	660
Pred. No.:	0	660.00	660
Score:	0	660.00	660
Percent Similarity:	0	100.00%	0
Best Local Similarity:	0	100.00%	0
Query Match:	0	100.00%	0
DB:	3		0
Length:	0		660
Matches:	0		660
Conservative:	0		0
Mismatches:	0		0
Indels:	0		0
Gaps:	0		0

US-09-851-410A-6 COPY 5147 7129 (1-1983) X AAB24120 (1-660)

Qy	1	ATGGCCCTCGGCCTATTGTTGTGCTCTCTCATGTTTTTGCCTATGCTGCGCCGCCA	60
Db	1	MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20
Qy	61	CCGCCCGGTACAGCTCTGGCCCGCTCGTGGCGCGCGCAGCGGGGTTCGGCGCGTGT	120
Db	21	ProProGlyGlnProSerGlyArgArgArgGlyAArgSerGlyGlySerGlyGlyGly	40
Qy	121	TTCTGGGGTGACCGGGTTGATTCTCAGCCCTTTCGCAATCCCTATATTCAATCAACAAAC	180
Db	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyIleHisProThrAsn	60
Qy	181	CCCTTCGCCCCCGATGTACCGCTCGCGCGGGGTGGACCTCTGTTTCGCCAACCCCGCC	240
Db	61	ProPheAlaProAspValThrAlaAlaGlyAlaGlyProArgValArgGlnProAla	80
Qy	241	CGACCACTCGGCTCCGCTTGGCGTCACAGGCCCCAGCGCCCGCGCTTCACCTCGTCT	300
Db	81	ArgProLeuGlySerAlaIleArgAspGlnAlaGlnArgProAlaValAlaSerArgArg	100
Qy	301	AGACCTACACAGCTGGGGCCGCGCGTAAACCGCGGTCTGCCTCCGCGCCCATGACACCCCG	360
Db	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro	120
Qy	361	CCAGTGCCTGATGTCGACTCCCGCGCGCCATCTTTGCGCCCGGCGAGTATTAACCTATCAACA	420

QY 721 GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATCTGTAACCAAGGCTGCGCTCC 780
 Db |||||
 QY 241 AlaserGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrPargSer 260
 Db |||||
 QY 781 GTCGAGACCTCTCGGGTGGCTGAGAGAGGCTACCTCTCGTCTTGTATGCTTTGCATA 840
 Db |||||
 QY 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
 Db |||||
 QY 841 CATGGCTCACTCGTAATCTCTATACATACACCTATACCGGTGCGCTCGGGCTGTTG 900
 Db |||||
 QY 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
 Db |||||
 QY 901 GACTTTGCGCTTCAGCTTGAGTTTCGCAACTTACCCCGGTAAACCAATACGCGGTC 960
 Db |||||
 QY 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
 Db |||||
 QY 961 TCCCGTTATTCACGACTGCTCGCCACCGCTTTCGTCGCGGTGCGGACGGAGTCCGAG 1020
 Db |||||
 QY 321 SerArgTyrSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
 Db |||||
 QY 1021 CTCACACACGCTGCTACCGCTTATCAAGGACCTCTATTTACTAGTACTATGCT 1080
 Db |||||
 QY 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
 Db |||||
 QY 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCCCTGTTCACCTTGTCTGACACTCTGCTT 1140
 Db |||||
 QY 361 ValGlyGluLeuGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
 Db |||||
 QY 1141 GCGCGCTGCGACAGAAATTCATTCTCGCGGTGCGCCAGCTGTTCTACTCCCGTCC 1200
 Db |||||
 QY 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
 Db |||||
 QY 1201 GTTGTCTCAGCAATGCGGACCGACTTAACTGTATACATCTGAGAGATGCTCAG 1260
 Db |||||
 QY 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
 Db |||||
 QY 1261 CAGGATAAGGATATGCAATCCCGCATGACATTCGCGGAGAAATCTCGTGTGTTATT 1320
 Db |||||
 QY 421 GlnAspLysGlyIleAlaIleProHisIleAspLeuGlyGluSerArgValValIle 440
 Db |||||
 QY 1321 CAGGATATGATAACCAACATGAACAAGATCGCGCGAGCTTCTCAGCCCACTCCGCG 1380
 Db |||||
 QY 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
 Db |||||
 QY 1381 CCTTCTCTCTCTCAGCTTAATGATGCTTGGCTCTCTCTCAGCGTCCGAGTAT 1440
 Db |||||
 QY 461 ProPheSerValLeuArgAlaAsnAspValLeuIlePheSerLeuLeuAlaGluTyr 480
 Db |||||
 QY 1441 GACAGTCCACTTATGGCTCTTCGACTGGCCAGTCTTATGTTTCTGACTCTGTGACCTG 1500
 Db |||||
 QY 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
 Db |||||
 QY 1501 GTTAATGTTGGACCGCGCGAGCGCTTGGCCGCTCGCTCGATTGGACCAAGGTACA 1560
 Db |||||
 QY 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLysValThr 520
 Db |||||
 QY 1561 CTTGACGCTGCGCCCTCTCCACATCCAGCAGTCTCGAAGACTTCTTGTCTCGCTCCG 1620
 Db |||||
 QY 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
 Db |||||
 QY 1621 CTCGCGGTAGCTCTCTTTCTGGAGGCGAGGACAACTAAAGCGGGTACCCCTATAAT 1680
 Db |||||
 QY 541 LeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
 Db |||||
 QY 1681 TATAACCACTAGTACGACCACTGCTTGTGAGATGCGCGCGGACCGGCTGCT 1740
 Db |||||
 QY 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
 Db |||||
 QY 1741 ATTTCCACTTACCACTAGCTGGTCTGGTCTGGTCTCCATCTCTGCGTGGCGCTT 1800
 Db |||||
 QY 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
 Db |||||
 QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGTTGAGGATACCTTGGACTACCTCCGCGCGCC 1860
 Db |||||

Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
 QY 1861 CATACTTTTGAATTTCTGCCAGAGTGCAGGCTTGGCTTCCAGGCTGCTTTC 1920
 Db |||||
 QY 621 HisThrPheAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
 Db |||||
 QY 1921 CACTCTACTGCTGCTGAGCTTTCAGCGCTTAAAGTGAAGTGGCTAAACTCGGAGTTG 1980
 Db |||||
 QY 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660
 Db |||||

RESULT 9
 AAO15699
 ID AAO15699 standard; protein; 660 AA.
 XX AAO15699;
 XX 08-NOV-2002 (first entry)
 XX Hepatitis E virus (Burma strain) ORF2 protein.
 DE HEV; enterically transmitted nonA/nonB hepatitis viral agent;
 KW Burma strain; bile; ORF2.
 XX Hepatitis E virus.
 OS US6379891-B1.
 PN 30-APR-2002.
 XX 19-APR-2000; 2000US-00553427.
 XX 17-JUN-1988; 88US-00208997.
 PR 11-APR-1989; 89US-00336672.
 PR 16-JUN-1989; 89US-00367486.
 PR 13-OCT-1989; 89US-00420921.
 PR 05-APR-1990; 90US-00505888.
 PR 15-APR-1991; 91US-00681078.
 PR 25-JUL-1994; 94US-00279823.
 PR 07-JUN-1995; 95US-00478507.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 XX Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam A, Fry KE;
 WPI; 2002-517277/55.
 DR N-PSDB; AAL50386.
 XX Detecting the presence of enterically transmitted nonA/nonB hepatitis viral (HEV) agents in bile samples from infected humans and monkeys using polymerase chain reaction.
 XX Disclosure; Col 19-24; 61pp; English.
 PS The invention comprises a method for detecting the presence of enterically transmitted nonA/nonB hepatitis viral (HEV) agents in a sample and isolating HEV agents or nucleic acid fragments produced by the agent. The method utilises PCR - using bile from a human or cynomolgus monkey actively infected with HEV as a source of the agent. The method of the invention is used for detecting the presence of a viral agent in a sample of cultured cells infected with the agent and isolating enterically transmitted nonA/nonB HEV agents or nucleic acid fragments produced by the agent. The present amino acid sequence represents the protein encoded by open reading frame 2 (ORF2) of a DNA sequence isolated from a Burmese strain of the Hepatitis E virus

XX SQ Sequence 660 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 660
 Score: 660.00 Matches: 660
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0
US-09-851-410a-6_COPY_5147_7129 (1-1983) x AA015699 (1-660)			
QY	1	ATGCGCCCTCGCGCCATTGTTGCTGCTCCATGTTTTCCTATGCTGCGCGCCA	60
DB	1	MetArgProArgProIleuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20
QY	61	CGCCCGCTCAGCCCTGCGCGCGCTGCTGGCGCGCGCAGCGCGGTTCGCGCGTGT	120
DB	21	ProProGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlySerGlyGlyGly	40
QY	121	TTCTGGGTGACCGGTTGATCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC	180
DB	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn	60
QY	181	CCCTTCGCCGCCGATGTACCGCTCGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC	240
DB	61	ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla	80
QY	241	CGACCACTCGGCTCGCTGGCGTGCACGAGGCCCGCGCGCGCTGCCTCAGCTGT	300
DB	81	ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg	100
QY	301	AGACCTACACAGCTGGGCGCGCGCTAAACCGCGGTGCTCCGCGCCATCACACCCGC	360
DB	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro	120
QY	361	CCAGTGCCTGATGTCGACTCCCGCGCGCCATCTTCGCGCGCAGTATAACCTCAACA	420
DB	121	ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr	140
QY	421	TCTCCCTTACTCTTCGCTGCGCGCAGGCACTAACCTGCTTCTTATGCGCGCCCTCT	480
DB	141	SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu	160
QY	481	AGTCGGCTTTACCCCTTCAGACGGCCACCATACCATATATGCGCCAGCAAGCTTCT	540
DB	161	SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180
QY	541	AATTATGCCAGTACCGGCTGCCGTCGCAAAATCCGTTACCGCGCGTGTCCCAAT	600
DB	181	AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn	200
QY	601	GCTGTCGCGGTACGCCATCTCCATCTCATTCTGCGCCACAGACCAACACCGCGAG	660
DB	201	AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr	220
QY	661	TCCGTTGATATCAATTCAATACCTCGAGGATGTTGTTATTTAGTCCAGCGCGGCATA	720
DB	221	SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle	240
QY	721	GCCTCTGAGCTGTGATCCCAAGTAGCGCCCTACACTATCGTAACCAAGCTCGCGCTCC	780
DB	241	AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer	260
QY	781	GTCGAGACCTCGGGTGCGCTGAGGAGGAGCTACTCTGCTGCTGTTATGCTTTGCATA	840
DB	261	ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280
QY	841	CATGCTCACTCGTAAATTCCTATACATACACCTATACCGGTGCCCTCGGCTGTGTG	900
DB	281	HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu	300
QY	901	GACTTTGCCCTTGAGCTTCAGTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGGTC	960
DB	301	AspPheAlaLeuLeuLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal	320
QY	961	TCCCGTTATCCAGCACTCGCCACCGCCCTTCGTCGCGGTCCGAGACGGAGCTCCGAG	1020
DB	321	SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu	340
QY	1021	CTCACCACACGGCTGTACCGCTTTATGAGGACCTCTATTTACTAGTACTAATGCT	1080
DB	341	LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly	360
QY	1081	GTCCGTGAGATCGCGCGGGATAGCCCTCACCTGTGTTCAACCTTGCTGACACTCTGCTT	1140
DB	361	ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu	380
QY	1141	GCGCGCTCGCCACAGAAATTGATTTGCTGCGCTGTGGCCAGCTGTTTACTCCCGTCCC	1200
DB	381	GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlnLeuPheTyrSerArgPro	400
QY	1201	GTTGCTCAGCAATGCGGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG	1260
DB	401	ValValSerAlaAsnGlyGluProThrValIlyLeuTyrThrSerValGluAsnAlaGln	420
QY	1261	CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGTTATT	1320
DB	421	GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle	440
QY	1321	CAGGATTATGATAACCAATGAAACAGATCGCGCGCGCTTCTCCAGCCCATCGCGC	1380
DB	441	GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg	460
QY	1381	CTTTTCTCTGCTCTCGAGCTAAATGATGTGCTTTCGCTCTCTCACCGCTCCCGAGTAT	1440
DB	461	ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr	480
QY	1441	GACCACTCACCTTATGGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG	1500
DB	481	AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu	500
QY	1501	GTTATGTTGCGACCGCGCGCGCGCTGCGCGCTGCTCGATGAGACCAAGTCCACA	1560
DB	501	ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr	520
QY	1561	CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTGCTCGCG	1620
DB	521	LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro	540
QY	1621	CTCCGCGGTAAAGTCTCTTTCTGGGAGCGAGCACAACTAAAGCGGGGTACCTTATAT	1680
DB	541	LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn	560
QY	1681	TATACACACTGCTAGGACCAACTGCTTGCAGAAATGCGCGCGGACCGGTCGCT	1740
DB	561	TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla	580
QY	1741	ATTTCCACTTACACCACTAGCTGCTGCTGCTGCTCCCTCTCCATTTCTGCGGTTGCCGT	1800
DB	581	IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal	600
QY	1801	TTAGCCCCCACTCTGCGCTAGCAATTGCTGAGGATACCTTGGACTACCTGCCCGCGCC	1860
DB	601	LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla	620
QY	1861	CATACCTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCCTTCAGGGCTGCGCTTC	1920
DB	621	HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe	640
QY	1921	CAGTCTACTGCTGCTGAGCTTCAGCGCTTAAAGTGAAGGTGGTAAACTCGGAGGTG	1980
DB	641	GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu	660
RESULT 10			
ADD24376			
ID	ADD24376	standard; protein; 660 AA.	
XX			
AC	ADD24376;		
DT	15-JAN-2004	(first entry)	
XX			

DE Hepatitis E virus strain Burma ORF 2.
XX non-A/non-B viral hepatitis; ET-NANB; HEV; vaccine.
XX Hepatitis E virus; strain Burma.
OS US2003124510-A1.
XX 03-JUL-2003.
XX 07-MAY-2001; 2001US-00851410.
XX 17-JUN-1988; 88US-00208997.
XX 11-APR-1989; 89US-00336672.
XX 16-JUN-1989; 89US-00367486.
XX 13-OCT-1989; 89US-00420921.
XX 05-APR-1990; 90US-00505888.
XX 05-APR-1991; 91US-00681078.
XX 25-JUL-1994; 94US-00279823.
XX 03-AUG-1998; 98US-00128275.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Reyes GR, Yarbough PO, Bradley DW, Krawczynski KZ, Tam AW;
PI Fry KE;
XX WPI; 2003-810978/76.
XX N-PSDB; ADD24374.
XX New proteins derived from an enterically transmitted non-A/non-B (ET-NANB) viral hepatitis agent, useful for diagnosing ET-NANB infection and producing vaccines for immunizing an individual against ET-NANB hepatitis viral agent.
XX Disclosure; SEQ ID NO 8; 8lpp; English.
XX The invention relates to a protein derived from an enterically transmitted non-A/non-B viral hepatitis agent. The viral proteins are useful for detecting infection by the enterically transmitted non-A/non-B (ET-NANB/HEV) hepatitis viral agent. The proteins are also useful for producing vaccines for immunising an individual against HEV hepatitis viral agent. The DNA fragments are useful for identifying and sequencing the entire viral agent and for assaying the viral agent in an infected sample. The present sequence is used in the exemplification of the invention.
XX Sequence 660 AA;
SQ
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x ADD24376 (1-660)
QY 1 ATGCGCCCTCGGCTATTTTGTCTCTCTCATGTTTGTCTATGCTGCCGGCCCA 60
DB 1 MetArgProArgProLeuLeuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGGTGAGCGCTGCGCCGCGCTGCGGCGCGCAGCGCGGTTCCGGCGGTGT 120
DB 21 ProProGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlySerGlyGlyGly 40
QY 121 TTCTGGGGTACCGGGTTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaAlaProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCCGCTGACCGCTGCGCGCGGCTGGACCTGCTGTCGCCAACCGGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyValGlyProArgValArgGlnProAla 80

QY 241 CGACCACTCGGCTCGCTTGGCGTGACCGGCCAGCGCCCGCGCTGCTCACGTCGT 300
DB 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTACCACTCGGCGCGCGCTAACCGCGGTCTCGCGGCCCATAGCACCCCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTCGACTCCCGCGGCGCATCTTGGCGCGGAGTATACCTATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTrpAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCCGTGGCCACCGGCACATACTGGTCTCTTTATGCGCCCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuLeuAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCCCTTCCAGGCGGCACCAATACCATATATATGCGCACCGAGCTTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGTTGCGCGTGCACAACTCGGTACCGCGCGTGGTCCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTGCGGGTTACGCCATCTCCATCTCATCTCTGCGCCACAGACCAACCCCGGACG 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
QY 661 TCGGTTGATATGAATCAATACTCGAGGATGTTTGTATTTAGTCCAGCCCGCAT 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGATCCCAAGTGAAGCGCTACACTATCGTAAACCAAGCTCGCGTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTCGGCTGGCTGAGGAGGCTACCTCTGCTCTGTATGCTTTGGCATA 840
DB 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGTCACTCGTAATTCCTATATAATACACCTATACCGGTGCCCTCGGGCTGTG 900
DB 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTCAGCTTCAGTTCCCACTTACCCCGGTAAACCAATACCGGTC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTTCCAGCAGCTGCTCGCCACCGCTTTCGTGCGGTGCGGACGGAGTCCGCG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACACGGCTGCTACCGCTTTATGAAGAGACTCTATTTTACTAGTACTATGGT 1080
DB 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTCAACCTGCTGACACTCTGCTT 1140
DB 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCCGACAGAAATTGATTTCGTGCGGTGGTGGCGAGCTGTTTCTACTCCCGTCCC 1200
DB 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGGCGAGCGGCTGTTAAGTTGTATACATCTCTAGAGATGCTCAG 1260
DB 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGTATTGCAATCCGCATGACATTGACCTCGGAGAACTCTCGTGGTGTATT 1320
DB 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGlnSerArgValIle 440

QY 1321 CAGGATTATGATAACCAACATGAACAGATCGCGAGCGCTCTCCAGCGCCCATCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTCTCTGCTTCGAGCTAATGATGCTTGGCTCTCTCTCACCGCTGCCAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTGCTTATGGCTCTCGACTGCGCCAGTTTATGTTCTGACTCTGTGACCTTG 1500
Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValLeu 500
QY 1501 GTTATGTTGGACCGCGCGCAGCGCTGTCGCCGTGCTCGATTGGACCAAGTCA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaAArgSerLeuAspTrpThrIysValThr 520
QY 1561 CTTGACGTGCGCCCTCTCCACATCCAGCAGTACTCGAAGACCTCTTTGTCCTGCGC 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPheValLeuPro 540
QY 1621 CTCGCGGTAAAGCTCTCTCTGGGAGGAGCAGCAACTAAAGCGGGTACCTTATAT 1680
Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGTAGCGACCAACTGCTTGTGAGAATGCCCGCGCACCGGTCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTCCACTTACACACTAGCTGGGTGCTGCTCCGCTCTCAATTCGTGGGTGCCGTT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTGCTTGAGGATACCTTGACTACCTGCCCGCGC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGATGATTCTGCCAGAGTCCGCCCTTGGCCTTTCAGGGCTGCGCTTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGAGCTTCAGCGCTTAAGATGAAGGTGGTAAACTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 11

AAW76369 standard; protein; 660 AA.

AAW76369;

03-DEC-1998 (first entry)

Hepatitis E virus hollow particle protein #2.

Hollow particle protein; virus; antibody; detection; immunoassay; infection.

Hepatitis virus.

JP10234383-A.

08-SEP-1998.

28-FEB-1997; 97JP-00062445.

28-FEB-1997; 97JP-00062445.

(DENK-) DENKA SEIKEN KK.

(KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.

WPI; 1998-535037/46.

N-PSDB; AAW61688.

PT Hepatitis E virus hollow particle poly(peptide(s) and nucleic acids
PT encoding it - useful for more accurate detection of HEV in samples, using
PT immuno-assays and nucleic acid hybridisation.

XX Claim 13; Page 24-26; 29pp; Japanese.

XX This sequence represents a Hepatitis E viral hollow particle protein.
CC This polypeptide can be used to raise antibodies to detect HEV infection
CC in samples, e.g. by immuno-assay based techniques, and the nucleic acid
CC can be used for the same in nucleic acid hybridisation assays. The
CC polypeptides and nucleic acids allow more accurate detection of HEV than
CC previously possible

XX Sequence 660 AA;

Alignment Scores:

Pred. No.: 0 Length: 660
Score: 559.00 Matches: 659
Percent Similarity: 99.85% Conservative: 0
Best Local Similarity: 99.85% Mismatches: 1
Query Match: 84.70% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAW76369 (1-660)

QY 1 ATGGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTTCCTATGCTGCTCGCGCGCA 60

Db 1 MetArgProArgProIleLeuLeuLeuLeuLeuMetPheLeuProMeLeuProAlaPro 20

QY 61 CCGCCCGCTCAGCGCTCGCGCGCGCTGCGCGCGCGCGCGCGCGGTTCGCGCGGTGGT 120

Db 21 ProProGlyGlnProSerGlyArgArgArgGlyArgArgSerGlyGlyGlyGly 40

QY 121 TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 180

Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60

QY 181 CCTTTCGCCCGCGATGTCCACCGCTGCGCGCGGGGTGGACCTCGTGTTCGCGCAACCGCGC 240

Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80

QY 241 CGACCACTCGGCTCGCTTGGCTGGCGACCGACCGCGCGCGCGCGCGCTTCACGTGCT 300

Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100

QY 301 AGACCTACACAGCTGGGGCGCGCGCTAACCGCGGTGCTCGCGCGCGCGCGCGCGCG 360

Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120

QY 361 CCAGTGCCTGATGTCGACTCCCGCGCGCGCGCTTTCGCGCGCGCGCGCTATACCTATCAACA 420

Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140

QY 421 TCTCCCTTACTCTTCGCTGGCCACCGCGCACTAACCTGGTTCTTTATGCGCGCGCTCTT 480

Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160

QY 481 AGTCGGCTTTACCCCTTCAGGACGGCACCAATCCCATATAATAGGCCACGAGAGCTTCT 540

Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180

QY 541 AATTATGCCAGTACCGGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600

Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200

QY 601 GCTGTGCGCGGTATAGCCCATCTCCATCTCATTTCTGGCCACAGACACACACCGCGGACG 660

Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220

QY 661 TCCGTTGATATGAATTCAATTAACCTCGACGAGTGTTCGTATTTAGTCCACCGCGGATA 720

Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240

QY 721 GCCTCTGAGCTTGTGATCCAAAGTGAAGCGCTACACTATCGTAACCAAGCGTGGCGCTCC 780
 Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyIrpArgSer 260
 QY 781 GTCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTTGTATGCTTTCGATA 840
 Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
 QY 841 CATGGCTCACTCGTAATTCCTATACCTATACACCTATACCGGTCCCTCGGGCTGTG 900
 Db 281 HasGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeu 300
 QY 901 GACTTTGCCCTTTCAGCTTCGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGCGGTC 960
 Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
 QY 961 TCCGTTATTCAGACTGCTGCCACCGGCTTCGTCGGGTGGCGGACCGGACTGCCGAG 1020
 Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
 QY 1021 CTCACCAACACCGCTCTACCCGCTTTATCAAGGACCTCTATTTTACTAGTACTAATGTT 1080
 Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
 QY 1081 GTCGTTGAGATCGGCGCGGATAGCCCTCACCTCTTCAACCTTCCTGACACTCTGCTT 1140
 Db 361 ValGlyGluIleGlyArgGlyAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
 QY 1141 GCGGCTCGGACAGAAATGATTCTGTCGGTGGTGGCGAGCTGTCTACTCCCGTCCC 1200
 Db 381 GlyGlyLeuProThrGluLeuLeuSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
 QY 1201 GTTGTCTCAGCAATCGGCGAGCGACTGTTAGTCTGTATACATCTCTAGAGATGCTCAG 1260
 Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
 QY 1261 CAGGATAAGGGTATTGCAATCCGCGATGACATTCGACCTCGGAGAACTCTGTGGGTATT 1320
 Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
 QY 1321 CAGGATTATGATAACCAATCAAGATCGGCCGACGCTTCTCAGCCCATCGCC 1380
 Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
 QY 1381 CTTTCTCTGCTCTCGAGCTAATGATGTCTTGGCTCTCTCACCCTGCGCGAGTAT 1440
 Db 461 PropheSerValLeuArgAlaAsnAspValLeuThrPheSerLeuThrAlaAlaGluTyr 480
 QY 1441 GACCACTCCACTATGGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
 Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
 QY 1501 GTTATGTTGCGACCGCGCGAGCGCTTGGCGGTCTGCTCGATTGGACCAAGTCA 1560
 Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspThrThrLysValThr 520
 QY 1561 CTTGACGGTGGCGCTCTCCACCATCCAGAGTACTCGAAGACCTTCTTTCCTCCGCG 1620
 Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPheValLeuPro 540
 QY 1621 CTCGCGGTAAGCTCTTTTTCGGGAGGAGGAGGACCAACTAAAGCCGGTACCGCTTATAAT 1680
 Db 541 LeuArgGlyLysLeuSerPheThrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
 QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAATGCCGCGGACCGGTCGCT 1740
 Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
 QY 1741 ATTTCACCTATACCACTAGCTGGTGGTGTGCTCCGCTCTCCATTTCTCGGGTTCGGT 1800
 Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
 QY 1801 TTAGCCCCCACTCTCGGCTAGCATTTGCTTGAGATACCTTGGAACCTACCTCGCCCGCC 1860

Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
 QY 1861 CATACTTTTATGATGTTCTGCCAGAGTCCGCCCTTGGCTTCAGGCTCGCTTTC 1920
 Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGluGlnGlyCysAlaPhe 640
 QY 1921 CAGTCTACTGTCGCTCAGCTTCAGGCGCTTAAGATGAAGTGGGTAAAACTCGGAGTTG 1980
 Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 12
 AAW71210
 ID AAW71210 standard; protein; 660 AA.
 XX AAW71210;
 XX 25-MAR-2003 (revised)
 DT 30-OCT-1998 (first entry)
 XX
 DE Protein encoded by ORF 2 of the Burmese isolate of ET-NANB.
 XX
 KW Enterically transmitted nonA/nonB hepatitis virus; identification; HEV;
 KW ET-NANB; detection; vaccine.
 XX
 OS Hepatitis virus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 496 /note= "not specified"
 FT
 XX US5789559-A.
 XX
 PD 04-AUG-1998.
 XX
 PF 25-JUL-1994; 94US-00279823.
 XX
 PR 17-JUN-1988; 88US-00208997.
 PR 11-APR-1989; 89US-00336672.
 PR 19-JUN-1989; 89US-00367486.
 PR 13-OCT-1989; 89US-00420921.
 PR 05-APR-1990; 90US-00505888.
 PR 05-APR-1991; 91US-00681078.
 XX
 (GENE-) GENELABS TECHNOLOGIES INC.
 XX
 PI Bradley DW, Reyes GR, Krawczynski KZ, Tam A, Fry KE, Yarbough PO;
 XX
 WPI; 1998-446186/38.
 XX
 PT Hepatitis E virus DNA - useful for e.g. virus detection and viral protein
 PT production.
 XX
 PS Disclosure; Col 63-68; 45pp; English.
 XX
 CC AAW71209-11 represent the proteins encoded by the open reading frames
 CC (ORFs) of the DNA sequence of the Burmese isolate of an enterically
 CC transmitted nonA/nonB viral hepatitis agent (ET-NANB). The nucleic acid
 CC sequence may be used for identifying and sequencing the entire viral
 CC agent (also referred to as HEV), detecting ET-NANB in infected samples,
 CC e.g. by specific amplification of virus-derived DNA sequences and for
 CC producing recombinant viral proteins for use in vaccines. (Updated on 25-
 CC MAR-2003 to correct PF field.)
 XX
 SQ Sequence 660 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 660
 Score: 559.00 Matches: 659
 Percent Similarity: 99.85% Conservative: 0
 Best Local Similarity: 99.85% Mismatches: 1
 Query Match: 84.70% Indels: 0
 DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAW71210 (1-660)

QY 1 ATGGCGCCTCGGCTATTTGCTGCTCTCTATGTTTTCCTATGCTGCGCGGCCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCGCGGTACGCGGTCTGGCGCGCGGTCTGGCGCGCGCGCGGTCTGGCGGTGGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlySerGlyGly 40
QY 121 TTCTGGGGTGACCGGGTGTATTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCCGATGTACCGGTGCGCGCGGGCTGACCTGTGTTCGCAACCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGCGTCTCGGTGACCGGTGACCGCGCGCGCGCGGTGCTCAAGTCGT 300
Db 81 ArgProLeuGlySerAlaIleTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACCTACCAAGTGGCGCGCGCGCTAACCGCGGTTCGCTCGCGCGCCATGACACCGCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAAGTCTGATGCTGACTCGCGCGCGCGCATCTTGGCGCGCGAGTATACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTCCGTGGCGCACCGCACCTAACCTGTTCTTATGCGCGCTCTT 480
Db 141 SerProLeuThrSerSerValAlaIleThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCCCTTCAGGACGGACCAATACCATATATATGGCGCGAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTCGCCGTGCACAAATCCGTATACCGCGCTGCTGCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaAargAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGCGGTACGCGCATCTCCATCTCATTCGGCGCACAGACACACACCGCGACG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATTAACCTCAGCGGATGTCGTATTTTTCAGTCCAGCGCGGATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTACGCGGTACACTATGTAACCAAGGCTGCGGTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgIleHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGGTGTGAGAGAGGCTACCTGCTGTGTTTATGCTTTTGATA 840
Db 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAAATTCCTACTATAATACACCTATACCGGTGCCCTCGGGCTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCTTGAGTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACCGGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATCCAGCACTGCTCGCACCGCTTCTGTCGGGTGCGGACGGACTCGCGAG 1020
Db 321 SerArgTyrSerSerThrAlaAargHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGGTGCTACCGGTCTTATGAAGGACCTCTATTTTACTAGTACTAATGGT 1080

Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCCGTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTGTGCTGACACTCTGCTT 1140
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GGGCGCTCGCGACAGAAATGATTTCGTGCGCTGGTGGCCAGCTGTTCTACTCCCGTCCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGCGGAGCCGACTGTTAAGTTGTATATCATCTGTAGAGAAATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGGTATTSCAATCCCGCATGACATTCACCTCGGAGAAATCTCGTGGTTATT 1320
Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
QY 1321 CAGGATTATGATAACCAACATGAACAGATCGGCGCGACTTCTCCAGCGCCCATCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTCTCTGCTTCGAGCTAATGATGCTGTTGGCTCTCTCTCATCCGCTGCGCGAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCAGTCCACTTATGCTCTTCGACTGCGCGCAGTTTATGTTTCTGACTCTGACCTTG 1500
Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSer***SerValThrLeu 500
QY 1501 GTTAATCTGCGACGGCGCGCGCTTGGCGCGTTCGATTCGATTCGACCAAGGTCA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTGACGGTCCGCCCTCTCCACATCCAGCAGTACTCGAAGACTTCTTTGCTCGCGC 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGGCGGTAAAGCTCTCTTCTGGAGCGGAGGACCAACTAAAGCGGGTACCTTATAAT 1680
Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTTCGAGAATCGCGCGGCGCACCGGCTGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACCACTAGCCTGGGTGCTGGTCCCTCTCCATTTCTCGGGTGGCGTT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACCTGCGCTAGCATTTGCTTGAGGATACCTTGGACTACCTGCGCGGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTCATGATTTCTGCCAGAGTGGCGCGCTTGGCTTACAGGCTGGCGTTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGAGCTTCAGCGCTTAAAGTGAAGTGGGTAAACTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 13

AAR96091
ID AAR96091 standard; protein; 549 AA.
XX
AC AAR96091;
XX
DT 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX
DE Hepatitis E virus (Burma strain) 62K antigen.
XX

KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
KW capsid.
XX
OS Hepatitis E virus; Burma strain.
XX
PN WO9612807-A2.
XX
PD 02-MAY-1996.
XX
PF 23-OCT-1995; 95WO-US013703.
XX
PR 24-OCT-1994; 94US-00327952.
XX
PR 13-OCT-1995; 95US-00542634.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
XX
DR N-PSDB; AAT27109.
XX
XX Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
PT diagnostic reagents for determining HEV infection and in vaccines.
XX
PS Claim 3; Page 83-84; 125pp; English.
XX
CC 62K antigen (AAR96091) comprises the C-terminal 549 amino acids of
CC hepatitis E virus (HEV) Burma strain capsid protein (see also AAR96089).
CC It is obt'd. by PCR amplification of HEV Burma ORF-2 (AAT27107), insertion
CC of amplified DNA (AAT27109) into vector pGEX and expression in E. coli
CC cells. Expression of full-length ORF-2 in Sf9 insect cells using a
CC baculovirus system also results in prodn. of 62K antigen (see also
CC AAR96101 and AAR96103). 62K represents an improved antigen, in comparison
CC to bacterial expressed proteins, for use in HEV diagnostic assays, and
CC also has excellent immunogenic properties. (Updated on 16-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 549 AA;
XX
Alignment Scores:
Pred. No.: 0 Length: 549
Score: 549.00 Matches: 549
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.18% Indels: 0
DB: 2 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAR96091 (1-549)
QY 334 GCGGTGCGTCCGGCCCATGACACCCGCGAGTGTGATGTGCTCCCGGGCGCCATC 393
DB 1 AlaValAlaProAlaHisAspThrProProValProAspValAspSerArgGlyAlaIle 20
QY 394 TTGCGCGCGAGTATAACCTATCAATCTCCCTTACCTCTCCGTGGCCACCGGCACT 453
DB 21 LeuArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaIleThrGlyThr 40
QY 454 AACCTGGTCTTTATGCGCGCCCTTTAGTCGGCTTTTACCCCTTCAGACGGACCAAT 513
DB 41 AsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACCCATATAATGGCCACCGAAGCTTCTAATATGCGCAGTACCGGGTTGCCGTCGCCACA 573
DB 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThr 80
QY 574 ATCCGTTACCGCCCGCTGGTCCCAATGCTGTCGCGGTTACGCCATCTCCATCTCATTC 633
DB 81 IleArgTyrArgProLeuValProAlaValAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCCACAGACACACACCGCGAGCCGCTGTGATGCAATTCATACCTCGACGGAT 693
DB 101 TrpProGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120

QY 694 GTTCGTAATTTAGTCCAGCCCGCATAGCCTCTGAGCTGTGTCATCCCAAGTGAGCGCTTA 753
DB 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTATCGTAACCAAGGCTGCGCTCCGTCCAGACCTCTGGGGTGGCTGAGGAGAGGCT 813
DB 141 HisTyrArgAsnGlnGlyTipArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCTGCTCTTGTATGCTTTGCATACATAGCTCAGTCTGTAATTCCTACTACTATAACA 873
DB 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180
QY 874 CCCTATACCGGTGCGCTCGGCTGTGGACTTTGCTTGGCTTGGCTTGGCTTGGCTTGGCT 933
DB 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200
QY 934 ACCCGCGGTAAACCAATACCGGCTCTCCGTTATTCAGACATGCTGCGCACCGCTT 993
DB 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
QY 994 CGTCGCGGTGCGGAGGACTGCGGAGCTCACCACCGGCTGCTACCGCTTTATGAAG 1053
DB 221 ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetIys 240
QY 1054 GACCTCTATTTTACTAGTACTAATGCTGCTGAGATCGGCGCGGATAGCCCTCACC 1113
DB 241 AspLeuTyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThr 260
QY 1114 CTGTTCAACCTGTGTGACACTCTGCTTGGCGGCTGCGCAGACAGATTGATTTGTCGGCT 1173
DB 261 LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAla 280
QY 1174 GGTGCGCAGCTGTTCTACTCCGCTCCCTGCTCTCAGCAATGGCGAGCCGCTTGAAG 1233
DB 281 GlyGlyGlnLeuPheTyrSerArgProValSerAlaAsnGlyGluProThrValIys 300
QY 1234 TTGTATACATCTGTAGAGATGCTCAGCAGATGAGGTATTGCAATCCCGCATCACATT 1293
DB 301 LeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIle 320
QY 1294 GACCTCGGAGAAATCTGTTGTTGTTATTCAGATTATGATTAACCAATGACAGATCGG 1353
DB 321 AspLeuGlyGluSerArgValIleGlnAspTyrAspAsnGlnHisGluGlnAspArg 340
QY 1354 CCGAGCGCTTCTCCAGCCCATCGCGCTTCTCTGCTCTCTGAGCTAAATGATGCTT 1413
DB 341 ProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu 360
QY 1414 TGGCTCTCTCTCACCGCTGCGGATGATGACAGTCCACTTATGCTTTCGACTGGCCCA 1473
DB 361 TrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPro 380
QY 1474 GTTTATGTTCTGACTCTGTGACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1533
DB 381 ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400
QY 1534 CGGTGCTCGATTGGACCAAGTCACTTCAGCGTCCGCCCTCTCCACCATCCAGCAG 1593
DB 401 ArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420
QY 1594 TACTCGAAGACCTCTTTGCTGCTGCGCTCGCGGTAAAGTCTCTTTCTGGGAGGACGC 1653
DB 421 TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGly 440
QY 1654 ACACTAAGCGGGTACCTTTATTAATTAACACCTAGCTAGGACCACTGCTTGTTC 1713
DB 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuVal 460
QY 1714 GAGATGCGCGCGGCGGCTGCTATTTCCACTTACACCATGCTGCTGGTCTGGT 1773
DB 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGly 480

QY 1774 CCCGTCCTCCATTTCTGGGTTGGCGTTTGTAGCCCCCAGCTCTGGCTAGCATTCCTTGAG 1833
 |||||||
 Db 481 ProValSerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGlu 500
 |||||||
 QY 1834 GATACCTTGAGACTACCTGGCCGCGCCATACATTTTGTATGATTCTGCCAGAGTGGCCG 1893
 |||||||
 Db 501 AspThrLeuAspTyrProAlaArgAlaHisThrPheAspAspPheCysProGluCysArg 520
 |||||||
 QY 1894 CCCCTGGCCTTCAGGGCTGGCTTTCAGTCTACTCTCGCTGAGCTTCAGCGCCTTAAG 1953
 |||||||
 Db 521 ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLys 540
 |||||||
 QY 1954 ATGAAGCTGGGTAAACTCGGGAGTTG 1980
 |||||||
 Db 541 MetLysValGlyLysThrArgGluLeu 549
 |||||||

RESULT 14
 AAE12980
 ID AAE12980 standard; protein; 660 AA.
 XX
 AC AAE12980;
 XX
 DT 28-JAN-2002 (first entry)
 XX
 DE Hepatitis E virus ORF2 encoded long capsid protein.
 XX
 KW Hepatitis E virus; HEV; capsid protein; antiviral; vaccine; therapy;
 KW detection; PORF2.
 XX
 OS Hepatitis E virus.
 XX
 FH Key Location/Qualifiers
 FT Region 221..660
 FT /note= "This region contains P1-P51 overlapping synthetic
 FT peptides"

XX WO200177156-A2.
 XX
 XX 18-OCT-2001.
 XX
 PF 03-APR-2001; 2001WO-US010696.
 XX
 PR 07-APR-2000; 2000US-0195380P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Fields HA, Khudyakov YE, Meng J;
 XX
 DR WPI; 2001-657163/75.
 XX
 XX C-terminal portion of the PORF2 protein of hepatitis E virus polypeptide
 PT containing a neutralizing epitope, useful to detect, prevent and treat
 PT hepatitis E virus infections.
 XX
 XX Example; Page 58-61; 62pp; English.
 PS
 CC The invention relates to a neutralising hepatitis E virus (HEV)
 CC polypeptide comprising C-terminal of HEV PORF2 protein containing at
 CC least one neutralising epitope. HEV is a non-enveloped virus. HEV
 CC polypeptide is useful for studying the pathogenesis of HEV and monitoring
 CC treatment efficacy in subjects undergoing treatment for HEV. The
 CC polypeptide is used as a reagent to detect HEV and as a vaccine to treat
 CC or prevent HEV infection. The present sequence is HEV long capsid protein
 CC encoded by open reading frame 2, also referred as PORF2
 XX
 SQ Sequence 660 AA;

Alignment Scores:
 Pred. No.: 0 Length: 660
 Score: 541.00 Matches: 641
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 81.97% Indels: 0

DB: 4 Gaps: 0
 US-09-851-410a-6_COPY_5147_7129 (1-1983) x AAE12980 (1-660)

QY 55 GCGCCACCGCCCGGTTCAGCGCTTCGGCCCGCTGCTGGCGGCGGCGACGCGCGGTTCGCGC 114
 |||||||
 Db 19 AlaProProGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlySerGly 38
 |||||||
 QY 115 GGTGGTTCTTGGGGTGACCGGGTGTATTCTCAGCCCTTCGCAATCCCTATATTCATCCA 174
 |||||||
 Db 39 GlyGlyPheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisPro 58
 |||||||
 QY 175 ACCAACCCCTTCGCCCGCCGATGTCAACCGCTGCGGCGCGGGCTGACCTCGTTTCGCAA 234
 |||||||
 Db 59 ThrAsnProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGln 78
 |||||||
 QY 235 CCGCCCGCAGCACTCGGCTCGCTTGGCTGTGACCGCCAGCCCGCGGCTTCCTCA 294
 |||||||
 Db 79 ProAlaArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSer 98
 |||||||
 QY 295 CGTCGTAGACTACCAAGCTGGGGCGCGCCCTAACCGGGTCTCGCTCGGCCCATGAC 354
 |||||||
 Db 99 ArgArgArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAsp 118
 |||||||
 QY 355 ACCCCGCGAGTGTGATGTGACTCCCGCGCGGCCCATCTTGGCGCGGCGAGTATAACTA 414
 |||||||
 Db 119 ThrProValProAspValAspSerArgGlyAlaAlaLeuArgArgGlnTyrAsnLeu 138
 |||||||
 QY 415 TCAACATCTCCCTTACCTCTTCGGTGGCCACCGGCACCTAACCTGGTCTTTATGCGGCC 474
 |||||||
 Db 139 SerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAla 158
 |||||||
 QY 475 CCTCTTAGTCGGCTTTTACCCCTTCAGGAGCGGCACCAATACCCATATAATGCGCCAGAA 534
 |||||||
 Db 159 ProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGlu 178
 |||||||
 QY 535 GCTTCTAATATATGCCAGTAGTACCGGGTTGCGCGTGCACAAATCCGTTTACCGCCGCTGTC 594
 |||||||
 Db 179 AlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuVal 198
 |||||||
 QY 595 CCCAATGCTGTCGGCGGTTCAGCCATCTCCATCTCATTTCTGGCCACAGACACACACC 654
 |||||||
 Db 199 ProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThr 218
 |||||||
 QY 655 CCGACGTCGGTTCATATGAATTCAAATACCTCGACGAGTCTTCGTATTTTAGTCCAGCC 714
 |||||||
 Db 219 ProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnPro 238
 |||||||
 QY 715 GGCATAGCCTCTGAGCTTGTGATCCCAAGTGAGCGCCTACACTATCTGTAACCAAGGCTGG 774
 |||||||
 Db 239 GlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyr 258
 |||||||
 QY 775 CGTCCGTCGAGACCTCTGGGGTGGCTGAGGAGAGCTACCTCTGCTGTGTTATGCTT 834
 |||||||
 Db 259 ArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeu 278
 |||||||
 QY 835 TGCATACATCGCTCACTCGTAAATTCCTATATACACCTATACCTGCTGCTGCTGGG 894
 |||||||
 Db 279 CysIleHisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGly 298
 |||||||
 QY 895 CTGTGGACCTTTCGCCCTTGAGCTTGTGATTCGCAACCTTACCCCGGTAAACCAATACG 954
 |||||||
 Db 299 LeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThr 318
 |||||||
 QY 955 CGGGTCTCCCGTTATTCCAGCACTGTGCGCCACCGCTTCGTGCGGTGCGGACGGGACT 1014
 |||||||
 Db 319 ArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThr 338
 |||||||
 QY 1015 GCGGAGCTACACACCGGCTGCTACCGCTTTATGAAGACCTCTATTATTACTAGTACT 1074
 |||||||
 Db 339 AlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThr 358
 |||||||
 QY 1075 AATGGTGTGCGTGAGATCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGCTGACACT 1134
 |||||||

Db 359 AsnGlyValGlyGluLeuGlyAlaLeuThrLeuPheAsnLeuAlaSerThr 378
 QY 1135 CTGCTGGCGGCTCCGACAGAAATGATTCTCGTGGTGGCCAGCTGTCTACTCC 1194
 Db 379 LeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyCysGlnLeuPheTyrSer 398
 QY 1195 CGTCCGCTGCTCAGCAATGGCGAGCGCTTAAGTTAGTTATACATCTCTAGAGAT 1254
 Db 399 ArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsn 418
 QY 1255 GCTCAGCAGGATAGGGTATTGCAATCCGCATGACATTGACCTCGGAGAATCTCGTGTG 1314
 Db 419 AlaGlnGlnAspLysGlyLeuAlaIleProHisAspIleAspLeuGlyGluSerArgVal 438
 QY 1315 GTTATTACAGATTATGATAACCAACATGAACAGATCGGCCGACGCTTCTCCAGCCCA 1374
 Db 439 ValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaPro 458
 QY 1375 TCGCGCCCTTCTCTGCTCTCGAGCTAATGATGTGCTTGGCTCTCTCACCCTGCC 1434
 Db 459 SerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAla 478
 QY 1435 GAGTATGACCACTCCACTATGGCTCTTCGACTGGCCAGTTATGTTCTGACTCTGTG 1494
 Db 479 GluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerVal 498
 QY 1495 ACCTTGGTTATGTTGCGACCGCGCGAGCGCTGCTCGCTCGATGGACCAAG 1554
 Db 499 ThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLys 518
 QY 1555 GTCACACTTGACGGTGGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTC 1614
 Db 519 ValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheVal 538
 QY 1615 CTGCGCTCCGCGGTAAGCTCTCTTCTGGGAGGCGAGCACAACTAAAGCCGGTACCT 1674
 Db 539 LeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrPro 558
 QY 1675 TATATTATACACCACTGCTAGCCACCACTGCTGTGCGAATGCCCGCGGACCGG 1734
 Db 559 TyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArg 578
 QY 1735 GTCGCTATTCTCACCCTACCACTAGCTGGGTGCTGCTCCGCTCTCCATTTCTCGGTT 1794
 Db 579 ValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSerAlaVal 598
 QY 1795 GCGGTTTGTAGCCCCCACTCTGCGTAGCATTTGAGGATACCTTGAGTACCTGCTGCC 1854
 Db 599 AlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAla 618
 QY 1855 CGCGCCCATCTTTTGTATGATTCTGCGCAGAGTGGCGCCCTTGGCTTCAGGCTGC 1914
 Db 619 ArgAlaHisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCys 638
 QY 1915 GCTTTCAGCTCTACTGCTGAGCTTCCAGCGCTTAAAGATGAAGGTGGGTAAACTCGG 1974
 Db 639 AlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArg 658
 QY 1975 GAGTTG 1980
 Db 659 GluLeu 660

RESULT 15

ID AAR96101
 AC AAR96101 standard; protein; 540 AA.
 XX
 DT 16-OCT-2003 (revised)
 DT 06-AUG-1996 (first entry)
 XX
 DE Hepatitis E virus (Burma strain) recombinant 62K antigen.

XX
 KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
 KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
 KW capsid.
 XX
 OS Hepatitis E virus; Burma strain.
 XX
 PN WO9612807-A2.
 XX
 PD 02-MAY-1996.
 XX
 PF 23-OCT-1995; 95WO-US013703.
 XX
 PR 24-OCT-1994; 94US-00327952.
 PR 13-OCT-1995; 95US-00542634.
 XX
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 XX
 PI Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
 XX
 WI 1996-230608/23.
 XX
 PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
 PT diagnostic reagents for determining HEV infection and in vaccines.
 XX
 PS Claim 5; Page 92-93; 125pp; English.
 XX
 CC DNA (AAT27109) coding for the 62K antigen (AAR96091) of hepatitis E virus
 CC (HEV) Burma strain capsid protein was cloned into baculovirus expression
 CC vector pBacII and recombinant 62K was expressed in Sf9 insect cells.
 CC High levels of expression were obtained, and the recombinant 62K was obtained in
 CC over 95% purity. However, C-terminal processing resulted in the deletion
 CC of 9 or 23 amino acids from 62K, giving 2 related polypeptide species
 CC (AAR96101 and AAR96103). Similar results were obtained with HEV Mexico 62K
 CC antigen (see also AAR96102 and AAR96104). Recombinant 62K represents an
 CC improved antigen, in comparison to bacterial expressed proteins, for use
 CC in HEV diagnostic assays, and also has excellent immunogenic properties.
 CC (Updated on 16-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 540 AA;

Alignment Scores:
 Pred. No.: 0 Length: 540
 Score: 540.00 Matches: 540
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 81.82% Indels: 0
 DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x AAR96101 (1-540)

QY 334 GGGTTCGCTCCGGCCCATGACACCCGCGCTGATGCTGACTCCGCGGCGCCATC 393
 Db 1 AlaValAlaProAlaHisAspThrProValProAspValAspSerArgGlyAlaIle 20
 QY 394 TTGCGCCGCGCAGTATACCTATCACTCTCCCTTACTCTTCCTGCGCCACCGGCACT 453
 Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThr 40
 QY 454 AACCTGGTTCCTTATGCGCCCTCTTAGTCGCTTTTACCTTACCTTCAGGACGGCACT 513
 Db 41 AsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn 60
 QY 514 ACCCATATAATGCGCACGGAAGCTCTTAATATATGCCAGTACCGGGTTGCCGCTGCCACA 573
 Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaAlaThr 80
 QY 574 ATCCGTTACCGCGCTGTCCTCCCAATGCTGTGCGCGGTAGCCCATCTTCCATCTCATTC 633
 Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
 QY 634 TGGCCACAGACCCACCACCCGCGCTGCTGATATGAATTCATTAACCTTCGCGGAT 693

101 TrpProGlnThrThrThrThrThrThrSerValAspMetAsnSerIleThrSerThrAsp 120
Qy GTTCGTATTTAGTCCAGCCCGGATAGCCCTGAGCTTGATGCCAAGTGAGCGCTA 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
Qy 754 CACTATCGTAACCAAGGTCGGCTCCGCTCCAGACCTCTGGGGTGGCTGAGGAGGAGCT 813
Db 141 HisTyrArgAsnGlnGlyTyrPargSerValGluThrSerGlyValAlaGluGluAla 160
Qy 814 ACCTCTGCTCTGTATGCTTTGTCATACATCGCTACCTCGTAATTCCTATACATAACA 873
Db 161 ThrSerGlyLeuValMetCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180
Qy 874 CCCTATACCGGTGCCCTCGGCTCTGTGGACTTGCCTTGAGCTTGAGTTTCGCAACCTT 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsnLeu 200
Qy 934 ACCCGCGGTAAACCAATACCGGGTCTCCCGTTATTCAGCACTGCTCGCCACCGCCTT 993
Db 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
Qy 994 CGTCGCGGTGCGGACGAGCTCCGAGCTCACACACGGTGTCTACCCGCTTTATGAAG 1053
Db 221 ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLys 240
Qy 1054 GACCTCTATTTTACTAGTACTAATGGTGTGGTGAGATCGCGCGGATAGCCCTCACC 1113
Db 241 AspLeuTyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThr 260
Qy 1114 CTGTTCAACCTTGCTGACACTCTGCTTGGCGGCTTCCGCGAGTATGATTTGCTCGGCT 1173
Db 261 LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAla 280
Qy 1174 GGTGGCCAGCTGTCTACTCCGCTCCCGTGTCTCAGCCCAATGGCGACCGCACTGTTAAG 1233
Db 281 GlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLys 300
Qy 1234 TTGTATACATCTGATAGAAATGCTCAGCAGGATAGGATAGGATATGCATCCCGCATGACATT 1293
Db 301 LeuTyrThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIle 320
Qy 1294 GACCTCGGAGAAATCTCGTGTGGTATTACAGATTATGATTAACCAACATGAACAAGATCGG 1353
Db 321 AspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArg 340
Qy 1354 CCGACGCTTCTCCAGCCCAATCGCGCCCTTCTCTGCTTCGAGCTAATGATGTGCTT 1413
Db 341 ProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu 360
Qy 1414 TGGCTCTCTCTCAGCTGCGGATGACAGTACAGTCCACTATGCTCTTCGACTGGCCCA 1473
Db 361 TrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerThrGlyPro 380
Qy 1474 GTTATGTTTCTGACTCTGTGACCTTGTGTTAATGTTCGACCGCGCGCAGCGCTTGCC 1533
Db 381 ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400
Qy 1534 CGGTGCTCGATTGACCAAGTCACTTACAGCTCGCCCTCTCCACCATCCAGCAG 1593
Db 401 ArgSerLeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420
Qy 1594 TACTCGAGACCTTCTTGTCTCGGCTCCGCGTAAGCTCTCTTCTGGGAGGCGAGGC 1653
Db 421 TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGly 440
Qy 1654 ACAACTAAAGCCGGGTACCTTTATATTATACACACTGCTAGCGACCACTGCTGCTGTC 1713
Db 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuVal 460
Qy 1714 GAGAAATGCCCGCGGACCGGTCGCTATTTCCACTTACACCACTAGCCTGGGTGCTGCT 1773
Db 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGly 480

Qy 1774 CCGCTCTCCATTTCTCGGGTTGCGGTTTGTAGCCCCCACTCTCGGCTAGCATTTGCTTGAG 1833
Db 481 ProValSerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGlu 500
Qy 1834 GATACCTTGGACTACCTCGCCGCGCCCATATCTTTTGATGATTTCTGCCAGAGTGCAGC 1893
Db 501 AspThrLeuAspTyrProAlaArgAlaHisThrPheAspAspPheCysProGluCysArg 520
Qy 1894 CCGCTTGGCCTTCAGGGCTGCGCTTCCAGTCTACTCTCGCTGAGCTTCAGCGCTTAAG 1953
Db 521 ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLys 540

Search completed: August 16, 2004, 13:47:46
Job time : 126.818 secs

Blank sheet

121 TTCTGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 CCCTTCGCCGCCGATGTACCGCTGGCGCGCGGCTGGACCTCGTGTTCGCCAACCGGCC 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 ProPheAlaProAspValThrAlaAlaAlaGlyProArgValArgGlnProAla 80
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 CGACCACTCGCTCCGCTGGCTGGCTGACAGGCCAGGCCCGCCCGCTGCTCAGTGTGT 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
81 ArgProLeuGlySerAlaTrpArgGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 AGACCTACCAACAGCTGGCGCGCGCTAACCGCGTGGCTCGCGGCCCATGACACCCCG 360
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 CCAGTGCCTGATCGACTCCCGCGCGCATCTTGGCGCGGAGTAACTATCAACA 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 TCTCCCTTACCTTCCGTTGGCGCACCGGCACTAACTGTGTTCTTATGCGCGCCCTCT 480
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 AGTCCGCTTTACCCCTTCAGCAGCGCACCAATACCATATATATGGCCACGGAAGCTTCT 540
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 AATTATGCCAGTACCGGGTGGCGTGGCCACCAATCCGTTACCGCGCGTGGTCCCAAT 600
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601 GCTGTCGGGGTTACGCCATCTCCATCTCATCTGTTGCGCCACAGACCAACCCCGGAG 660
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
661 TCCGTTGATATGAATCAATAACTCGACGATGTTGCTATTTTGTCCAGCCCGGCATA 720
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
721 GCCTCTGAGCTTGATGCCAAGTACGAGCGCTACACTATGTAACCAAGCTGGCGCTCC 780
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
781 GTCCAGACCTCTGGGTGGCTGAGGAGGCTACCTCTGCTTCTGTTGCTTGTTCGATA 840
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
841 CATGGCTCACTCGTAATTCCTATATACTAATACACCTATACCGGTGCGCTCGGGCTGTG 900
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
901 GACTTTGCCCTTGAGCTTGATTTGCAACCTTACCCCGGTAAACCAATACGCGGGTC 960
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
961 TCCCGTTATTCAGCACTCTCGCACCGCTTCGTCGGGTGCGGACGCGGACGCGGAG 1020
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1021 CTCACCAACACCGCTCTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGTT 1080
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1081 GTCCGTGATCGCGCGGATAGCCCTCACCTGTTCACTTCTGACACTCTGCTT 1140
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1141 GCGCGCTCGCGACAGATTGTTTCGTGGCTGGTGGCCAGCTGTCTTACTCCCGTCCC 1200
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlnLeuPheTyrSerArgPro 400
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1201 GTTGTCTCAGCCAATGCGCGCGCACTGTTTAAAGTTGTATATACATCTCTAGAGAATGCTCAG 1260

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1261 CAGGATAAGGGTATTCAATCCCGCATGACATTGACCTCGAGAAATCTCGTGGTTATT 1320
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 440
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1321 CAGGATTATGATTAACCAACATGAACAGATCGCGCCAGCGCTTCTCCAGCCCCATCGCGC 1380
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1381 CCTTCTCTCTCTCGAGCTAATGATGTCTTGTCTCTCTCTCACCCTGCGCGAGTAT 1440
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1441 GACCACTCCACTTATGGCTTCTCGACTGGCCCCAGTTATGTTTCTGACTCTGACCTTG 1500
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1501 GTTAATGTTGCGACCGCGCGGAGCGGTTGCGCGTCCGCTCGATTGGACCAAGTCA 1560
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1561 CTTGACGGTCGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTCTGCCG 1620
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1621 CTCGCGGTAAGCTCTCTTTCGGGAGGAGGAGGAGCAACATAAGCCGGGTACCTTATAAT 1680
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1681 TATAACCACTCTAGCAGCACTGCTTGTGCGAAGTCCCGCCGGGACCCGGGTGCT 1740
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
561 TyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1741 ATTTCACCTTACCACTAGCCTGCTGCTGGTCCGCTCTCCATTTCTGCGGTGCGGT 1800
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1801 TTAGCCCCCACTCTCGCTAGCATGCTTGAGGATACCTTGGACTACCTGCGCCGCGCC 1860
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1861 CATACCTTTGATCATTTCTGCCAGAGTCCCGCCCTTGGCTTCCAGGCTGCGCTTTC 1920
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1921 CAGTCTACTGCTCGCTCAGCTTACGCGCTTAAAGATGAAGTGGGTAAACCTCGGGAGTTG 1980
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 2

B44212

structural protein 2 precursor - hepatitis E virus (strain Mexico)

C:Species: hepatitis E virus

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999

C:Accession: B44212

R:Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.

A:Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HE)

A:Reference number: A44212, MUID:93079857, PMID:1448913

A:Accession: B44212

A:Molecule type: genomic RNA

A:Residues: 1-659 <HUA>

A:Cross-references: GB:M74506; NID:9330017; PIDN:AAA45732.1; PID:9330020

C:Superfamily: hepatitis E virus structural protein 2

C:Keywords: structural protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-659/Product: structural protein 2 #status predicted <SP2>

Alignment Scores:

Pred. No.: 2,05e-109

Length: 659

Score: 120.00

Matches:

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.18% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x B44212 (1-659)

QY 490 TTACCCCTTCAGGCGGACCAATACCCATATAATGGCCAGCAAGCTTCTAATTATGCC 549
DB 164 LeuProLeuGlnAspGlyThrAsnThrHisLeuMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGGTTCGCCGTCACCAATCCGTTACCCCGCGCTGGTCCCAATGCTCTCGGC 609
DB 184 GlnTyrArgValAlaArgAlaThrileArgTyrArgProLeuValProAsnAlaValGly 203
QY 610 GGTTACGCCATCTCCATCTCTGCGCCACAGACACCAACCCCGACGTCGGTTGAT 669
DB 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCATAAATCCCTCAGCGATGTTCTGATTTTAGTCCAGCCCGGATAGCCTCTGAG 729
DB 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 CTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGCTCCGTCGAGACC 789
DB 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThr 263
QY 790 TCTGGGTGGCTGAGGAGGAGCTACCTCTGCTTGTATGCTTTGCAATACATGGCTCA 849
DB 264 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283

RESULT 3

VHWHE

structural protein 1 - hepatitis E virus (strain Burma)

C:Species: hepatitis E virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 23-Jul-1999
C:Accession: B40778; A40236
R: Tam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.C.; Bradley, D.W.; Fry, K.E.; Reyes, G.
Virology 185, 120-131, 1991
A: Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length vi
A: Reference number: A40778; PMID:92024067; PMID:1926770
A: Accession: B40778
A: Molecule type: genomic RNA
A: Residues: 1-123 <TAM>
A: Cross-references: GB: M73218; NID: g330023; PIDN: AAA45735.1; PID: g330025
R: Ray, R.; Jameel, S.; Manivel, V.; Ray, R.
Virology 189, 359-362, 1992
A: Title: Indian hepatitis E virus shows a major deletion in the small open reading frame
A: Reference number: A40236; PMID:92295577; PMID:1534953
A: Accession: A40236
A: Molecule type: genomic RNA
A: Residues: 1-57, 'P', 59-102, 'P', 104-123 <RAY>
C: Superfamily: hepatitis E virus structural protein 1
C: Keywords: structural protein

Alignment Scores:
Pred. No.: 1 85e-98 Length: 123
Score: 109.00 Matches: 109
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.52% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x VHWHE (1-123)

QY 2 TCGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCAC 61
DB 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGCCGGTTCAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGT 121
DB 35 ArgProValSerArgLeuAlaValValGlyGlyAlaAlaValProAlaValVal 54

QY 122 TCTGGGTGACCGGTTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAACC 181
DB 55 SerGlyValThrGlyLeuIleLeuSerProSerGlnSerProIlePheIleGlnProThr 74
QY 182 CTTTCGCCCCGATGTCACCGCTGCGCGCGCGCTGGACCTGCTGTTGCTGCCAACCGGCC 241
DB 75 ProSerProProMetSerProLeuArgProGlyLeuAspLeuValPheAlaAsnProPro 94
QY 242 GACCACTCGGCTCGCTTGGCGTGACAGGCGCCGCGCGCGCGCGCTGCTCAGCTCGTA 301
DB 95 AspHisSerAlaProLeuGlyValThrArgProSerAlaProProLeuProHisValVal 114
QY 302 GACCTACCAACAGCTGGCGCGCGCGCGCGC 328
DB 115 AspLeuProGlnLeuGlyProArgArg 123

RESULT 4

C44212

structural protein 1 - hepatitis E virus (strain Mexico)

C:Species: hepatitis E virus

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999

C:Accession: C44212

R: Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.
Virology 191, 550-558, 1992

A: Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HE

A: Reference number: A44212; PMID:93079857; PMID:1448913

A: Accession: C44212

A: Molecule type: genomic RNA

A: Residues: 1-123 <HUA>

A: Cross-references: GB: M74506; NID: g330017; PIDN: AAA45731.1; PID: g330019

C: Superfamily: hepatitis E virus structural protein 1

C: Keywords: structural protein

Alignment Scores:
Pred. No.: 4.45e-51 Length: 123
Score: 61.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.24% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x C44212 (1-123)

QY 2 TCGCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCAC 61
DB 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGCCGGTTCAGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGT 121
DB 35 ArgProValSerArgLeuAlaValValGlyGlyAlaAlaValProAlaValVal 54
QY 122 TCTGGGTGACCGGTTGATTTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAACC 181
DB 55 SerGlyValThrGlyLeuIleLeuSerProSerGlnSerProIlePheIleGlnProThr 74

QY 182 CCT 184

DB 75 Pro 75

RESULT 5

B40236

structural protein 1 - hepatitis E virus (strain Indian)

C:Species: hepatitis E virus

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999

C:Accession: B40236

R: Ray, R.; Jameel, S.; Manivel, V.; Ray, R.

Virology 189, 359-362, 1992

A: Title: Indian hepatitis E virus shows a major deletion in the small open reading frame

A: Reference number: A40236; PMID:92295577; PMID:1534953

A: Accession: B40236

A: Molecule type: genomic RNA

A: Residues: 1-41 <RAY>

C: Superfamily: hepatitis E virus structural protein 1

C;Keywords: structural protein

Alignment Scores:
 Pred. No.: 0.000128 Length: 41
 Score: 14.00 Matches: 14
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.12% Indels: 0
 DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x B40236 (1-41)

QY 287 TTGCTCACGCTAGACCTACACAGCTGGGCGCGCGC 328
 |||||
 Db 28 LeuProHisValValAspLeuProGlnLeuGlyProArgArg 41

RESULT 6

B86242
 hypothetical protein [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: B86242
 R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B86242
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-337 <STO>
 A;Cross-references: GB:AE0051172; NID:g4874264; PIDN:AAD31329.1; GSPDB:GN00141
 C;Genetics:
 A;Map position: 1

Alignment Scores:
 Pred. No.: 8.28 Length: 337
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.39% Indels: 0
 DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x B86242 (1-337)

QY 1953 CTTAAGCGCTGAAGCTCAGCGACGT 1927
 |||||
 Db 149 LeuLysAlaLeuLysLeuSerAspSer 157

RESULT 7

T12281
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - Poecilia sphenops mitochondrion
 C;Species: mitochondrion Poecilia sphenops
 C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 03-Jun-2002
 C;Accession: T12281
 R;Breden, F.; Pracek, M.; Rashed, M.; Taphorn, D.; de Figueiredo, C.A.
 submitted to the EMBL Data Library, October 1997
 A;Description: Molecular phylogeny of a live-bearing fish genus Poecilia (Cyprinodontifc
 A;Reference number: Z17479

A;Accession: T12281
 A;Status: preliminary; translated from GB/EMBL/DDBU
 A;Molecule type: DNA
 A;Residues: 1-348 <BRE>
 A;Cross-references: EMBL:AF031390; NID:g2653584; PID:g2653585; PIDN:AAB87615.1
 C;Genetics:
 A;Genome: mitochondrion
 C;Superfamily: NADH dehydrogenase (ubiquinone) chain 2

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Alignment Scores:
 Pred. No.: 8.24 Length: 348
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x T12281 (1-348)

QY 1910 GCTGCGCTTCCAGTCTACTGTGCTG 1936
 |||||
 Db 281 AlaAlaLeuSerSerLeuLeuSerLeu 289

RESULT 8

H96635
 unknown protein, 73054-77165 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: H96635
 R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: H96635
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-840 <STO>
 A;Cross-references: GB:AE0051173; NID:g6751691; PIDN:AAF27674.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: T7P1.16
 A;Map position: 1

Alignment Scores:
 Pred. No.: 7.31 Length: 840
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x H96635 (1-840)

QY 1907 AGGCTGCGCTTCCAGTCTACTGTGCG 1933
 |||||
 Db 114 ArgAlaAlaLeuSerSerLeuLeuSer 122

RESULT 9

F83928
 hypothetical protein BH2230 [imported] - Bacillus halodurans (strain C-125)
 C;Species: Bacillus halodurans
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C;Accession: F83928
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: F83928
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1091 <STO>
 A;Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BA05949.1; GSPDB:GN00
 C;Genetics:
 A;Experimental source: strain C-125

A;Gene: BH2230

Alignment Scores:
Pred. No.: 7.06 Length: 1091
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x F83928 (1-1091)

QY 19 TTGTTGCTGCTCTCATGTTTTCCT 45

DB 12 LeuLeuLeuLeuLeuMetPheLeuPro 20

RESULT 10

H58208

protamine II - black rat snake (fragment)

C;Species: Elaphe obsoleta quadrivittata

C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 07-May-1999

C;Accession: H58208

R;Hunt, J.G.; Kasinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.

J. Biol. Chem. 271, 23547-23557, 1996

A;Title: Protamines of reptiles.

A;Reference number: A58208; MUID:96394458; PMID:8798564

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-41 <HUN>

C;Superfamily: sperm histone

Alignment Scores:

Pred. No.: 107 Length: 41
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x H58208 (1-41)

QY 76 TCTGGCCGCGTCGTGGGCGGCG 99

DB 13 SerGlyArgArgArgGlyArg 20

RESULT 11

S19708

hypothetical protein C - Herpetosiphon aurantiacus

C;Species: Herpetosiphon aurantiacus

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Jan-1995

C;Accession: S19708

R;Erdmann, D.; Duesterhoeft, A.; Kroeger, M.

Eur. J. Biochem. 202, 1247-1256, 1991

A;Title: Cloning and molecular characterization of the HgiCI restriction/ modification

A;Reference number: S19706; MUID:92111503; PMID:1662609

A;Accession: S19708

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-74 <KRO>

A;Cross-references: EMBL:X55138

Alignment Scores:

Pred. No.: 98.7 Length: 74
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x S19708 (1-74)

QY 827 ACAAGACCAGAGGTAGCCTCCTCC 804

Db 64 ThrArgProGluValAlaSerSer 71

RESULT 12

D81144

hypothetical protein NMB0903 [imported] - Neisseria meningitidis (strain MC58 serogroup C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C;Accession: D81144

R;Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755; PMID:10710307

A;Accession: D81144

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-97 <TET>

A;Cross-references: GB:AE002442; GB:AE002098; NID:G7226133; PIDN:AAF41311.1; PID:G722614

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB0903

Alignment Scores:

Pred. No.: 95.2 Length: 97
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x D81144 (1-97)

QY 837 GCAAAGCATACACAGCAGAGGT 814

DB 43 AlaLysHisAsnLysThrArgGly 50

RESULT 13

JT0593

hypothetical 15K protein (hgidiim 3' region) - Herpetosiphon aurantiacus

N;Alternate names: ORF15

C;Species: Herpetosiphon aurantiacus

C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 04-Mar-2000

C;Accession: JT0593; S21948

R;Duesterhoeft, A.; Kroeger, M.

Gene 106, 87-92, 1991

A;Title: Cloning, sequence and characterization of m5C-methyltransferase-encoding gene,

A;Reference number: JT0592; MUID:92039068; PMID:1937045

A;Accession: JT0593

A;Molecule type: DNA

A;Residues: 1-138 <DUE>

A;Cross-references: EMBL:X55141; NID:G48771; PIDN:CRA38940.1; PID:G48772

A;Experimental source: strain Hpa2

C;Superfamily: Herpetosiphon aurantiacus hypothetical 15K protein (hgidiim 3' region)

Alignment Scores:

Pred. No.: 90.7 Length: 138
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x JT0593 (1-138)

QY 822 ACCAGAGGTAGCCTCCTCCTCAGC 799

DB 43 ThrArgGlySerLeuLeuSer 50

RESULT 14

D72394

hypothetical protein TW0292 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: D72394
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: D72394
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-166 <ARN>
A;Cross-references: GB:AE001711; GB:AE000512; NID:g4980788; PIDN:AAD35380.1; PID:g498079
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TW0292

Alignment Scores:
Pred. No.: 88.5 Length: 166
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x D72394 (1-166)

QY 1728 GCACGGGTGCTATTTCCACTTA 1751

Db 19 AlaProGlyArgTyrPheHisLeu 26

RESULT 15

F87649

ExbD/TolR family protein [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C;Accession: F87649

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: F87649

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-172 <STO>

A;Cross-references: GB:AB005673; NID:g13424916; PIDN:AAK25194.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC3232

Alignment Scores:

Pred. No.: 88 Length: 172
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x F87649 (1-172)

QY 79 GGCGCGGTGTCGGCGCGCAGC 102

Db 16 GlyArgArgGlyArgGser 23

Search completed: August 16, 2004, 13:57:39
Job time : 40.2737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:36:13 ; Search time 18.2638 Seconds
(without alignments)
11307.070 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129

Perfect score: 660

Sequence: 1 ATGGCGCTCGGCTATTTT.....GTAAACTCGGAGTTGTAG 1983

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Word size: 1

Total number of hits satisfying chosen parameters: 283186

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO_spool_p/US09851410/runat_16082004_125716_9625/app_query.fasta_1.4238
-DB=SwissProt 42 -Qfmt=fastan -SUFFIX=oli.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410@cgn 1.1.43 @runat_16082004_125716_9625 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	660	100.0	660	1 VST2 HEVBU	P29326 hepatitis e
2	408	61.8	660	1 VST2 HEVPA	P33426 hepatitis e
3	128	19.4	660	1 VST2 HEVH	Q04611 hepatitis e
4	120	18.2	485	1 VST2 HEVHR	Q00270 hepatitis e
5	120	18.2	659	1 VST2 HEVME	Q03500 hepatitis e
6	109	16.5	123	1 VST1 HEVBU	P29325 hepatitis e
7	76	11.5	123	1 VST1 HEVHY	Q04612 hepatitis e
8	61	9.2	123	1 VST1 HEVME	Q03499 hepatitis e
9	9	1.4	337	1 IF33 ARATH	Q9C522 arabidopsis
10	8	1.2	121	1 YD2M HERAU	Q8xh7r clostridium
11	8	1.2	138	1 YD2M HERAU	P25279 herpetosiph
12	8	1.2	166	1 LED1 THEMA	Q9wyc8 thermotoga
13	8	1.2	199	1 PSBQ CHLRE	P12852 chlamydomon
14	8	1.2	202	1 PSBQ VOLCA	Q41643 volvox cart
15	8	1.2	218	1 LOLE XANCP	Q8pc65 xanthomonas
16	8	1.2	233	1 UL03 HSV2H	P28279 herpes simp
17	8	1.2	241	1 RR2 FORPU	P51249 porphyra pu
18	8	1.2	258	1 VSP3 TRIGA	Q13063 trimeresuru

19	8	1.2	273	1 DEFC ARATH	Q9fuz2 arabidopsis
20	8	1.2	277	1 THTR MYCLE	Q50036 mycobacteri
21	8	1.2	318	1 ATH4 ARATH	P92953 arabidopsis
22	8	1.2	331	1 MANI MOUSE	Q9wu40 mus musculu
23	8	1.2	358	1 PTPA HUMAN	Q15257 homo sapien
24	8	1.2	412	1 YNP1 CAEEL	P34554 caenorhabdi
25	8	1.2	507	1 MKR3 HUMAN	Q13064 homo sapien
26	8	1.2	511	1 HRDB STRCO	P18183 streptomyce
27	8	1.2	540	1 SNB2 HUMAN	Q13425 homo sapien
28	8	1.2	634	1 ICPG SYNY3	P37979 synechocyst
29	8	1.2	732	1 PSAB CYAME	Q85fy6 cyanidiosch
30	8	1.2	748	1 MEPA RAT	Q64230 rattus norv
31	8	1.2	789	1 ARNT HUMAN	P27540 homo sapien
32	8	1.2	790	1 ARNT RABIT	Q02748 oryctolagus
33	8	1.2	791	1 ARNT MOUSE	P53762 mus musculu
34	8	1.2	800	1 ARNT RAT	P41739 rattus norv
35	8	1.2	837	1 ATS4 HUMAN	O75173 homo sapien
36	8	1.2	860	1 MUTS PASMU	P57972 pasteurella
37	8	1.2	884	1 SYA XYLEFT	Q9ph22 xyliella fas
38	8	1.2	884	1 SYA XYLEFT	O87f43 xyliella fas
39	8	1.2	928	1 CHS2 EXODE	P30601 exophiala d
40	8	1.2	955	1 GCSP BRAJA	Q89186 bradyrhizob
41	8	1.2	993	1 TSH DROME	P22265 drosophila
42	8	1.2	1003	1 MBD6 HUMAN	Q96dn6 homo sapien
43	8	1.2	1164	1 CNA2 HUMAN	Q9h723 homo sapien
44	8	1.2	3969	1 HRX HUMAN	Q03164 homo sapien
45	7	1.1	39	1 PRT1 BUFJA	P24641 bufo japoni

ALIGNMENTS

RESULT 1

VST2 HEVBU
ID VST2 HEVBU STANDARD; PRT; 660 AA.

AC P29326;
DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 01-FEB-1994 (Rel. 28, Last annotation update)

DE Structural protein 2 precursor (ORF2).

OS Hepatitis E virus (strain Burma) (HEV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage;

OC Hepatitis E-like viruses.

OX NCBI_TaxID=31767;

RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=92024067; PubMed=1926770;

RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,

Fry K.E., Reyes G.R.;

RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the

full-length viral genome.";

RL Virology 185:120-131(1991).

CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING

THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA

BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See http://www.isb-sib.ch/announce/

or send an email to license@isb-sib.ch).

CC

CC EMBL; M73218; AAA45736.1; -

DR PIR; C40778; VHWWH2.

DR InterPro; IPR004261; SP2.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF03014; SP2; 1.

Signal.

KW Q1643 volvox cart

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 660 STRUCTURAL PROTEIN 2.

SQ SEQUENCE 660 AA; 70978 MW; 5832A013CC4A61C CRC64;

Alignment Scores:

Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST2_HEVBU (1-660)

```
QY 1 ATCCGCCCTCGGCTATTTTGTGCTCTCTCATGTTTGTGCTATGCTGCGCGGCCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20

QY 61 CCSCCGGTCAGCGCTGCGCGCGCTGCTGGGCGCGGAGCGCGGTTCCGGCGTGGT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgSergGlyGlySerGlyGly 40

QY 121 TTTCTGGGGTGACCGGGTGTATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
DB 41 PheTyrGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60

QY 181 CCCTTCGCCGCCGATGTCACCTGCGCGCGGCGGTGACCTCGTGTGCCAACCGGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyValArgValArgGlnProAla 80

QY 241 CGACCACTCGGCTCCGCTGGCGTGACCGCGCGCGCGCGCGCGCGCTGCTCAGCTCGT 300
DB 81 ArgProLeuGlySerAlaThrArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100

QY 301 AGACCTACCAACAGCTGGGCGCGCGCTAACCGCGTCCGTCGCGGCCCATGACACCCCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120

QY 361 CCGAGTCCGCTGATGTCAGTCCCGCGCGCCCATCTTGGCGCGGAGTAAACCTATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140

QY 421 TCTCCCTTACCTCTTCGCGGCGCACCGGCACTAACTGGTCTTTATGCGCGCCCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160

QY 481 AGTCCGCTTTTACCTCTCAGGACGGCACCAATACCATATATGCGCACCGGAGCTTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180

QY 541 AATTATGCCAGTACCGGGTTCGCGTCCACCAATCGTTTACCGCGCGTGGTCCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200

QY 601 GCTGTCGCGGTTACGCCATCTCCATCTCATCTGCGCACAGACCCACCCACCGCGACG 660
DB 201 AlaValGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrThrProThr 220

QY 661 TCGGTTGATGATGAATCAATAAAGTTCGACGAGATGTTCTGATATTTAGTCCAGCCGCGATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240

QY 721 GCCTCTGAGTTGTGATCCCAAGTAGCGGCTTACATATCGTAACCAAGCTCGCGCTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260

QY 781 GTCGAGACCTCTCGGGTGGCTGAGGAGGAGGCTACCTCTGCTGTGTTATGCTTTTCATA 840
DB 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280

QY 841 CATGGCTCACTCGTAAATTCCTACTATAATACACCTATACCGGTGCGCTCGGGCTGTG 900
DB 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300

QY 901 GACTTTGCCCTTACGCTTGTAGTTTCGCAACCTTACCCCGGTAAACCAATAACGCGGTC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
```

RESULT 2

VST2_HEVBU

ID VST2_HEVBU

AC P33426;

STANDARD;

PRT;

660 AA.

DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Structural protein 2 precursor (ORF2).
 OS Hepatitis E virus (strain Pakistan) (HEV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=33774;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92115700; PubMed=1731327;
 RA Tsarev S.A., Emerson S.D., Reyes G.R., Tsareva T.S., Legters L.J.,
 RA Malik I.A., Iqbal M., Purcell R.H.;
 RT "Characterization of a prototype strain of hepatitis E virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
 CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
 CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
 CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M80581; AAA45727.1; -;
 DR InterPro: IPR004261; SP2.
 DR InterPro: IPR008975; Viral_cap_coat.
 DR Pfam: PF03014; SP2; 1.
 KW SIGNAL.
 FT SIGNAL. 1 22 BY SIMILARITY.
 FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
 SQ SEQUENCE 660 AA; 70980 MW; 8085BC53CFB46FD3 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 660
 Score: 408.00 Matches: 508
 Percent Similarity: 99.80% Conservatives: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 61.82% Indels: 0
 DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST2_HEVPA (1-660)

QY 289 GCTCAGCTCGTAGACCTACACAGCTGGGGCGGCGCCCTAACCGCGTGGCTCGCGGCC 348
 DB 97 AlaSerArgArgProThrThrAlaGlyAlaAlaProLeuThrAlaValaProAla 116
 QY 349 CATGACACCCGCCAGTGGCTGATGCGACTCCCGCGGCGCATCTTGGCGCGGAGTAT 408
 DB 117 HisAspThrProProValProAspValAspSerArgGlyAlaAlaLeuArgGlnTyr 136
 QY 409 AACCTATCAACATCTCCCTTACCTTCCGTGGCCACCGGCACTAACCTGGTTCTTTAT 468
 DB 137 AsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyr 156
 QY 469 GCGGCGCTCTAGTCGGCTTTTACCCCTTCAGGACGGACCAATACCATATATATGGCC 528
 DB 157 AlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAla 176
 QY 529 ACGGAAGCTTCTAAATTATCCAGTACCGGGTTGCCGGTGCACAACTCCGTTACCGCCG 588
 DB 177 ThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgPro 196
 QY 589 CTGGTCCCCAATGCTGTCGGCGGTGTACGCCATCTCCATCTCATCTTGGCCACAGACCAAC 648
 DB 197 LeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThr 216
 QY 649 ACCACCCGACGTGGCTGATATGAATCAATAACCTCGACGGAGTTCGTAATTTAGTC 708
 DB 217 ThrThrProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuVal 236

QY 709 CAGCCGCGCATAGCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCTGTAACCAA 768
 DB 237 GlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGln 256
 QY 769 GGCTGGCGCTCGCTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGCTCTTGT 828
 DB 257 GlyTyrArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuVal 276
 QY 829 ATGCTTTGCATACATAGGCTCACTCGTAATTTCTTATCTACTATAACACCCCTATACCGGTGCC 888
 DB 277 MetLeuCyIleHisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAla 296
 QY 889 CTCGGCTCTTGGACTTTCCCTTGGCTTGTAGTTTTCGCAACTTACCCCGGTACACAC 948
 DB 297 LeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThr 316
 QY 949 AATACGGCGTCTCCGGTTATTCAGCACTGTCCGCCACCGCTTCGTCCGGTGGCGAC 1008
 DB 317 AsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAsp 336
 QY 1009 GGGACTGCCGAGCTCACACACGGCTGCTACCCGCTTTATGAGGACCTCTATTTACT 1068
 DB 337 GlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThr 356
 QY 1069 AGTACTAATGGTGTGGTGAGATCGCGCGGATAGCCCTCACCGCTGTCAACCTTGCT 1128
 DB 357 SerThrAsnGlyValGlyGluLeuGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAla 376
 QY 1129 GACACTCTGCTTGGCGGCTGCCGACAGAAATGATTTCGTCCGCTGGTGGCGAGCTGTC 1188
 DB 377 AspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPhe 396
 QY 1189 TACTCCGCTCCGTTGTCTCAGCCCATGGCGAGCGGACTGTAAAGTTGTATACATCTGTA 1248
 DB 397 TyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal 416
 QY 1249 GAGAAATGCTCAGCAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGATCT 1308
 DB 417 GluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSer 436
 QY 1309 CGTGTGGTTATTTCAGGATTATGATAACCAACATGAACAGATCGCGCGGCTCTTCCA 1368
 DB 437 ArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 456
 QY 1369 GCGCATCGCGCCCTTCTCTGCTCGAGCTAATGATGCTTGGCTCTCTCTCAC 1428
 DB 457 AlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThr 476
 QY 1429 GCTGCGGAGTATGACGAGTCCACTTATGGCTCTTCGACTGGCCGAGTTTATGTTCTGAC 1488
 DB 477 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 496
 QY 1489 TCTGTGACCTTGGTTAATGTTGCGACCGCGCGCGGCTTGGCGGCTCGCTCGATTGG 1548
 DB 497 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrp 516
 QY 1549 ACCAAGGTCACACTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTC 1608
 DB 517 ThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhe 536
 QY 1609 TTTGTCTGCGCGCTCCGCGGTAAGCTCTTTCTGCGAGGAGGACACAACTAAACCCGG 1668
 DB 537 PheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGly 556
 QY 1669 TACCCCTTAATTAATTAACACCACTGCTGACGACCACTGCTTGTGAGAAATGCCCGCGG 1728
 DB 557 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGly 576
 QY 1729 CACCGGGTGGCTTATTTCCACTTACACCACTAGCTGGGTGGTGGTCCGCTCTCCATTTCT 1788
 DB 577 HisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSer 596

```
QY 1789 GCGTTGCCGTTTACGCCCTCT 1815
DB 597 AlaValAlaValLeuAlaProHisSer 605
RESULT 3
VST2_HEVMY STANDARD; PRT; 660 AA.
ID VST2_HEVMY AC Q04611;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227573; PubMed=8470371;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
RL Virus Genes 7:95-109(1993).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10330; BAA01174.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW SIGNAL.
FT CHAIN 1 22 BY SIMILARITY.
FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70998 MW; 3A82A4EA255C6253 CRC64;
Alignment Scores:
Pred. No.: 1.3e-117 Length: 660
Score: 128.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.39% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST2_HEVMY (1-660)
QY 1 ATGCCGCTCGGCTATTTGCTGCTCTCATGTTTGTGCTATGCTGCCCGCCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGCTCAGCCGCTCGGCGCGCTGCGGCGGCGCAGCGCGGTTCGGCGGTGGT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlyGlyGly 40
QY 121 TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAACAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCCGATGTCAACGCTCGCGCGGCGTGACCTCGTGTTCGCAACCCGCC 240
DB 61 ProPheAlaProAspValThraAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTTGGCGTGAACAGGCCCGCCCGCGCTGCTCAGCTCGT 300
DB 1789 GCGTTGCCGTTTACGCCCTCT 1815
DB 597 AlaValAlaValLeuAlaProHisSer 605
RESULT 3
VST2_HEVMY STANDARD; PRT; 660 AA.
ID VST2_HEVMY AC Q04611;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227573; PubMed=8470371;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
RL Virus Genes 7:95-109(1993).
CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10330; BAA01174.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW SIGNAL.
FT CHAIN 1 22 BY SIMILARITY.
FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70998 MW; 3A82A4EA255C6253 CRC64;
Alignment Scores:
Pred. No.: 1.3e-117 Length: 660
Score: 128.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.39% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST2_HEVMY (1-660)
QY 1 ATGCCGCTCGGCTATTTGCTGCTCTCATGTTTGTGCTATGCTGCCCGCCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGCTCAGCCGCTCGGCGCGCTGCGGCGGCGCAGCGCGGTTCGGCGGTGGT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlyGlyGly 40
QY 121 TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCAAACAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCCGATGTCAACGCTCGCGCGGCGTGACCTCGTGTTCGCAACCCGCC 240
DB 61 ProPheAlaProAspValThraAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTTGGCGTGAACAGGCCCGCCCGCGCTGCTCAGCTCGT 300
DB 1789 GCGTTGCCGTTTACGCCCTCT 1815
DB 597 AlaValAlaValLeuAlaProHisSer 605
RESULT 3
VST2_HEVRH STANDARD; PRT; 485 AA.
ID VST2_HEVRH AC Q00270;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Structural protein 2 (Fragment).
OS Hepatitis E virus (isolate Rhesus) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31766;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K.,
RA Win K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90274; BAA20910.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
FT NON_TER 1 1
FT NON_TER 485 485
SQ SEQUENCE 485 AA; 52317 MW; 5A4FD9273AC74F94 CRC64;
Alignment Scores:
Pred. No.: 1.08e-109 Length: 485
Score: 120.00 Matches: 120
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.18% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST2_HEVRH (1-485)
QY 853 GTAATTCTCTATCTACTAATACACCTATACCGGTCCCTCGGCTGTGGACTTGCCTT 912
DB 153 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 172
QY 913 GAGCTTCAGTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGGTCTCCCGTTATTC 972
DB 173 GluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyrSer 192
QY 973 AGCACTCTCCCAACCCCTTCGTCGCGGTCCGACCGGAGCTGCCGAGCTCACCAACG 1032
DB 193 SerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGluLeuThrThr 212
QY 1033 GCTGCTACCCGCTTTATGAGGACCTCTATTACTAGTACTATGCTGCGGTGAGATC 1092
DB 213 AlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGlu 232
```

```
QY 1093 GGCGCGGGATAGCCCTCAACCTGTTCAACCTGTTGACACTCTGCTGGCGCGCTGCCG 1152
Db 233 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuPro 252
QY 1153 ACAGATTGATTTCGTGGCTGGTGGCCAGCTGTTCTACTCCCGTCCCGTGTCTCAGCC 1212
Db 253 ThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAla 272

RESULT 5
VST2_HEVME
ID VST2_HEVME STANDARD; PRT; 659 AA.
AC Q03500;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor.
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31768;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93079857; PubMed=1448913;
RA Huang C.-C., Nguyen D., Fernandez J., Yun K.-Y., Fry K.E.,
RA Bradley D.W., Tam A.W., Reyes G.R.;
RT "Molecular cloning and sequencing of the Mexico isolate of hepatitis
RT E virus (HEV).";
RL Virology 191:550-558 (1992).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M74506; AAA45732.1; -.
DR PIR; B44212; B44212.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW Signal.
FT SIGNAL. 1 22 BY SIMILARITY.
FT CHAIN 23 659 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 659 AA; 70640 MW; CF75E75EFD8FBE2C CRC64;

Alignment Scores:
Pred. No. 1 03e-109 Length: 120 Matches: 659
Score: 120.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 18.18%
Indels: 0
Gaps: 0
DB: 1

US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST2_HEVME (1-659)

QY 490 TTACCCCTTCAGAGCGGACCAATACCCATATAATGGCCACGAGCTTCTAATTATGCC 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGGTTCGGCGGCACATTCCTACCGCCCGCTGGTCCCAATCTGTCCGC 609
Db 184 GlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203
QY 610 GGTTAGCCCATCTCCATCTCATTTGGCCACAGACACACACCCAGCTCCGTTGAT 669
Db 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCAATAACCTCGACGGATGTCGTATTTTAGTCAGCCCGGCATAGCCTCTGAG 729
```

```
Db 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 CTTGTGATCCCAAGTGAAGCGCTACACTATCGTAACAAAGCTGGCGCTCGAGACC 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 263
QY 790 TCTGGGGTGGCTGAGGAGGAGGCTACCTGCTGGTCTTGTATGCTTTGCATACATGGCTCA 849
Db 264 SerGlyValAlaGluGlnGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283

RESULT 6
VST1_HEVBU
ID VST1_HEVBU STANDARD; PRT; 123 AA.
AC P29325;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 1 (ORF3).
OS Hepatitis E virus (strain Burma) (HEV), and
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31767, 33774;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Burma;
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Pakistan;
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563 (1992).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M73218; AAA45735.1; -.
DR EMBL; M60581; AAA45726.1; -.
DR PIR; B40778; VHWVHE.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
SQ SEQUENCE 123 AA; 12676 MW; 8A5A798B1B74EDE5 CRC64;

Alignment Scores:
Pred. No. 9.66e-99 Length: 123
Score: 109.00 Matches: 109
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 16.52%
Indels: 0
Gaps: 0
DB: 1

US-09-851-410A-6_COPY_5147_7129 (1-1983) x VST1_HEVBU (1-123)

QY 2 TCGGCCCTCGGCCTATTTTGTGCTGCTCCTCATGTTTTCCTATGTCGCCGCCGCAC 61
Db 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGCCCGGTTCAGCCGCTCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGGT 121
```


Pred. No.: 55-2 Length: 121
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.21% Indels: 0
 DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x RL7_CLOPE (1-121)

QY 83 GCCGTCGTGGCGCGCGAGCGCG 106

Db 38 AlaValValGlyGlyAlaAlaAa 45

RESULT 11

YD2M_HERAU
 ID YD2M_HERAU STANDARD; PRT; 138 AA.

AC P25279;

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 15.4 kDa protein in HgiDII M5 region (ORF15).

OS Herpetosiphon aurantiacus (Herpetosiphon giganteus).

OC Bacteria; Chloroflexi; Herpetosiphonales; Herpetosiphonaceae;

OC Herpetosiphon.

OC NCBI_TaxID=65;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HPA2;

RX MEDLINE=92039068; PubMed=1937045;

RA Duesterhoeft A., Kroeger M.;

RT "Cloning, sequence and characterization of m5C-methyltransferase-

RT encoding gene, hgiDII (GRCGAC), from Herpetosiphon giganteus strain

RT Hpa2.";

RL Gene 106:87-92(1991).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X55141; CAA38940.1; --

DR PIR; J0593; J0593.

KW Hypothetical protein; Restriction system; DNA-binding.

FT DNA_BIND 17 38 H-T-H MOTIF (POTENTIAL).

SQ SEQUENCE 138 AA; 15396 MW; E9CC70A0E11785A2 CRC64;

Alignment Scores:

Pred. No.: 54.1 Length: 138
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.23% Indels: 0
 DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x YD2M_HERAU (1-138)

QY 822 ACCAGAGGTAGCTCTCTCTCAGC 799

Db 43 ThrArgGlySerLeuLeuLeuSer 50

RESULT 12

LED1_THEME

ID LED1_THEME STANDARD; PRT; 166 AA.

AC Q3WYCB;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE 3-isopropylmalate dehydratase small subunit 1 (EC 4.2.1.33)

DE (Isopropylmalate isomerase 1) (Alpha-IPM isomerase 1) (IPMI 1).

GN LEUD1 OR TM0292.

OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
 NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
 RX MEDLINE=9287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
 Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima."
 RL Nature 399:323-329(1999).
 CC -|- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
 CC and 3-isopropylmalate, via the formation of 2-isopropylmaleate.
 CC -|- CATALYTIC ACTIVITY: 3-isopropylmaleate = 2-isopropylmaleate +
 CC H(2)O.
 CC -|- CATALYTIC ACTIVITY: 2-isopropylmaleate + H(2)O = 2-
 CC isopropylmaleate.
 CC -|- PATHWAY: Leucine biosynthesis; second step.
 CC -|- SUBUNIT: Heterodimer of leuC and leuD (By similarity).
 CC -|- SIMILARITY: Belongs to the leuD family. LeuD 2 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE001711; AAD35380.1; --
 DR PIR; D72394; D72394.
 DR TIGR; TM0292; --
 DR HAMAP; MF 01032; -; 1.
 DR InterPro; IPR000573; Aconitase_C.
 DR Pfam; PF00694; Aconitase_C; 1.
 KW Leucine biosynthesis; Lyase; Complete proteome.
 SQ SEQUENCE 166 AA; 18734 MW; D8ACF340205111A4 CRC64;

Alignment Scores:
 Pred. No.: 52.7 Length: 166
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.21% Indels: 0
 DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x LED1_THEME (1-166)

QY 1728 GCACCGGTCGCTATTTCACCTTA 1751
 |||||
 Db 19 AlaProGlyArgTyrPheHisLeu 26

RESULT 13

PSBQ_CHLRE
 ID PSBQ_CHLRE STANDARD; PRT; 199 AA.
 AC P12852;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Oxygen-evolving enhancer protein 3, chloroplast precursor (OE3).
 GN PSBQ.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OC NCBI_TaxID=3055;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=137C / CC-125;

RA Mayfield S.P., Schirmer-Rahire G., Frank H., Zuber H., Rochaix J.-D.;
RT "Analysis of the genes of the OEB1 and OEE3 proteins of the
RL Plant Mol. Biol. 12:683-693(1989).
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC with the photosystem II complex.
CC -1- SIMILARITY: Belongs to the psbQ family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC ENBL: X13832; CAA32061.1; --
DR PIR: S05509; S05509.
DR InterPro: IPR008797; PsbQ.
DR Pfam: PF05757; PsbQ; 1.
KW Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
KW Thylakoid; Membrane.
FT TRANSIT 1 51 CHLOROPLAST.
FT CHAIN 52 199 OXYGEN-EVOLVING ENHANCER PROTEIN 3.
SQ SEQUENCE 199 AA; 21824 MW; 8CCF077334CF79D6 CRC64;

Alignment Scores:
Pred. No.: 51.3 Length: 199
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1993) x PSBQ_CHLRE (1-199)

QY 277 CGCCCGCGTTCCTCAGTCGT 300
Db 10 ArgProAlaValAlaSerArg 17

RESULT 14
PSBQ_VOLCA STANDARD; PRT; 202 AA.
AC Q41643;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oxygen-evolving enhancer protein 3, chloroplast precursor (OEE3).
GN PsbQ.
OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=f. Nagariensis / HK10;
RX MEDLINE=96376499; PubMed=8781179;
RA Choi G., Przybylska M., Straus D.;
RT "Three abundant germ line-specific transcripts in Volvox carteri
RL encode photosynthetic proteins.";
RL Curr. Genet. 30:347-355(1996).
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC with the photosystem II complex.
CC -1- SIMILARITY: Belongs to the psbQ family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL: U22330; BAB40980.1; --
DR PIR: S72224; S72224.
DR InterPro: IPR008797; PsbQ.
DR InterPro: IPR006311; Tat.
DR Pfam: PF05757; PsbQ; 1.
DR TIGRFAMs: TIGR01409; TAT signal seq; 1.
KW Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
KW Thylakoid; Membrane.
FT TRANSIT 1 54 CHLOROPLAST (POTENTIAL).
FT CHAIN 55 202 OXYGEN-EVOLVING ENHANCER PROTEIN 3.
SQ SEQUENCE 202 AA; 22290 MW; 3103DDFC84CC3180 CRC64;

Alignment Scores:
Pred. No.: 51.2 Length: 202
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1993) x PSBQ_VOLCA (1-202)

QY 277 CGCCCGCGTTCCTCAGTCGT 300
Db 10 ArgProAlaValAlaSerArg 17

RESULT 15
LOLB_XANCP STANDARD; PRT; 218 AA.
AC Q8PC65;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer-membrane lipoprotein lolB precursor.
GN LOLB OR XCC0870.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Parah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.P.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RL host specificities.";
RL Nature 417:459-463(2002).
CC -1- FUNCTION: Plays a critical role in the incorporation of
CC lipoproteins in the outer membrane after they are released by the
CC lolA protein (By similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (By similarity).
CC -1- SIMILARITY: Belongs to the lolB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AE012186; AM40185.1; -.
DR HAMAP; MF_00233; -, 1.
DR InterPro; IPR004565; LolB.
DR InterPro; IPR000437; ProK_lipoprot_S.
DR Pfam; PF03550; LolB; 1.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Chapterone; Outer membrane; Lipoprotein; Transport; Protein transport;
KW Signal; Complete proteome; Palmitate.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 218 OUTER-MEMBRANE LIPOPROTEIN LolB.
FT LIPID 21 21 N-Palmitoyl cysteine (Potential).
FT LIPID 21 21 S-Diacylglycerol cysteine (Potential).
SQ SEQUENCE 218 AA; 23298 MW; 25A715B827CF782C CRC64;

Alignment Scores:
Pred. No.: 50.6 Length: 218
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x LolB_XANCP (1-218)

QY 326 GGCGCGCCCGCTGTGTAGGT 303
Db |||||||
30 GlyAlaAlaProAlaValValGly 37

Search completed: August 16, 2004, 13:48:43
Job time : 28.2638 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:39:23 ; Search time 97.0734 Seconds
(without alignments)
12890.708 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129

Perfect score: 660

Sequence: 1 ATGCGCCCTCGGCTATTTT.....GTAAACTCGGAGTTGTAG 1983

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2033934

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09851410/runat_16082004_125717_9638/app_query.fasta_1.4238
-DB=SPTRMBL_25 -QWMT=fastan -SUFFIX=oli.rspt -MINMATCH=0.1 -LOOPCL=0
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09851410 @CGN 1.1.218 @runat_16082004_125717_9638 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: sp archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvrius.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	456	69.1	660	12 Q9E8G5	Q9E8G5 hepatitis e

2	403	61.1	660	12 Q9WQA0	Q9WQA0 hepatitis e
3	396	60.0	660	12 Q89468	Q89468 hepatitis e
4	362	54.8	660	12 Q9W7W9	Q9W7W9 hepatitis e
5	344	52.1	344	12 Q9WJZ5	Q9WJZ5 hepatitis e
6	284	43.0	660	12 Q81871	Q81871 hepatitis e
7	282	42.7	660	12 Q9WLL4	Q9WLL4 hepatitis e
8	275	41.7	659	12 Q8JT32	Q8JT32 hepatitis e
9	264	40.0	660	12 Q89411	Q89411 hepatitis e
10	259	39.2	660	12 Q89419	Q89419 hepatitis e
11	251	38.0	660	12 Q91855	Q91855 hepatitis e
12	247	37.4	605	12 Q9WML0	Q9WML0 hepatitis e
13	214	32.4	660	12 Q91856	Q91856 hepatitis e
14	212	32.1	248	12 Q81860	Q81860 hepatitis e
15	197	29.8	660	12 Q81878	Q81878 hepatitis e
16	174	26.4	283	12 Q9YF81	Q9YF81 hepatitis e
17	168	25.5	168	12 Q9WJX0	Q9WJX0 hepatitis e
18	163	24.7	525	12 Q39947	Q39947 hepatitis e
19	149	22.6	149	12 Q9W9E8	Q9W9E8 hepatitis e
20	149	22.6	150	12 Q57042	Q57042 hepatitis e
21	149	22.6	660	12 Q88985	Q88985 hepatitis e
22	144	21.8	205	12 Q9QP98	Q9QP98 hepatitis e
23	133	20.2	135	12 Q91169	Q91169 hepatitis e
24	126	19.1	227	12 Q9W827	Q9W827 hepatitis e
25	125	18.9	258	12 Q81861	Q81861 hepatitis e
26	124	18.8	227	12 Q56048	Q56048 hepatitis e
27	124	18.8	227	12 Q56049	Q56049 hepatitis e
28	122	18.5	436	12 Q9W148	Q9W148 hepatitis e
29	119	18.0	137	12 Q8B6J4	Q8B6J4 hepatitis e
30	119	18.0	137	12 Q8B6J3	Q8B6J3 hepatitis e
31	119	18.0	137	12 Q8B6J2	Q8B6J2 hepatitis e
32	119	18.0	137	12 Q8B6J1	Q8B6J1 hepatitis e
33	119	18.0	137	12 Q8B6J0	Q8B6J0 hepatitis e
34	119	18.0	137	12 Q8B6I9	Q8B6I9 hepatitis e
35	119	18.0	137	12 Q8B6I8	Q8B6I8 hepatitis e
36	119	18.0	137	12 Q8B6I7	Q8B6I7 hepatitis e
37	119	18.0	137	12 Q8B6I6	Q8B6I6 hepatitis e
38	119	18.0	137	12 Q8B6I5	Q8B6I5 hepatitis e
39	119	18.0	137	12 Q8B6I4	Q8B6I4 hepatitis e
40	119	18.0	137	12 Q8B6I3	Q8B6I3 hepatitis e
41	119	18.0	137	12 Q8B6I2	Q8B6I2 hepatitis e
42	119	18.0	137	12 Q8B6I1	Q8B6I1 hepatitis e
43	119	18.0	137	12 Q8B6I0	Q8B6I0 hepatitis e
44	119	18.0	137	12 Q8B6H9	Q8B6H9 hepatitis e
45	119	18.0	137	12 Q8B6H8	Q8B6H8 hepatitis e

ALIGNMENTS

RESULT 1
Q9E8G5 PRELIMINARY; PRT; 660 AA.
ID Q9E8G5
AC Q9E8G5
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Abb-2B;
RX MEDLINE=20271579; PubMed=10813471;
RA van Cuyck-Gandre H., Zhang H.Y., Tsarev S.A., Warren R.L.,
RA Caudill J.D., Snellings N.J., Begot L., Innis B.L., Longer C.F.;
RT "Phylogenetically distinct hepatitis E viruses in Pakistan.";
RL Am. J. Trop. Med. Hyg. 62:187-189 (2000).
DR EMBL; AF185822; AAG16766.1;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.

SQ	SEQUENCE	660 AA; 70903 MW; 1F506BE3CFE3BACE CRC64;	
Alignment Scores:			
Pred. No.:	0	Length: 660	
Score:	456.00	Matches: 556	
Percent Similarity:	99.82%	Conservative: 0	
Best Local Similarity:	99.82%	Mismatches: 1	
Query Match:	69.09%	Indels: 0	
DB:	12	Gaps: 0	
US-09-851-410a-6_COPY_5147_7129 (1-1983) x Q9B8G5 (1-660)			
QY	310	ACAGCTGGGCGCGCGCTAACCGGCTCGCTCCGCGCCATGACACCCGCGCAGTGCCT	369
DB	104	ThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrProValPro	123
QY	370	GATGTGCGCTCCGCGCGCGCATCTTGGCGCGCGCAGTATAACCTATCAACATCTCCCTT	429
DB	124	AspValAspSerArgGlyAlaAlaLeuArgGlnTyrAsnLeuSerThrSerProLeu	143
QY	430	ACCTCTTCGCGCGCGCGCATTAACCTGCTTATGCGCGCGCTCTTATGCGCTT	489
DB	144	ThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeuSerProLeu	163
QY	490	TTACCCCTTCAGAGCGCACCAATACCATATATATGCGCGCGCGCATCTTAATATGCG	549
DB	164	LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla	183
QY	550	CAGTACCGGCTTGGCGCGCGCATCAATCCGTTACCGCGCGCTGGTCCCAATGCTGCGCG	609
DB	184	GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly	203
QY	610	GGTACGCCATCTCCATCTCATCTGCGCGCGCATCAACACACCGCGCGCGCTGAT	669
DB	204	GlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrThrThrThrThrThrThr	223
QY	670	ATGAATTCATACCTCGAGGATGCTGATTTTACGCGCGCGCGCATGCTGCTGAG	729
DB	224	MetAsnSerIleThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr	243
QY	730	CTTGATGATCCCAAGTGGCGCGCTACACTATCGTAACCAAGGCTGGCGCTCGCGAGACC	789
DB	244	LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThr	263
QY	790	TCTGGGTGGTGGAGGAGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	849
DB	264	SerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer	283
QY	850	CTCGTAATTCCTATCTATACACCTATACCGGCTGCGCGCGCTGCTGCTGCTGCTGCTG	909
DB	284	ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla	303
QY	910	CTTGAGCTTGAGTTTGGCAACTTACCCCGCGGTAAACCAATACCGCGGCTTCCCGTTAT	969
DB	304	LeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyr	323
QY	970	TCCAGCATCTGCGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1029
DB	324	SerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGluLeuThrThr	343
QY	1030	ACGGCTGCTACCGGCTTTTGAAGGACCTCTATTTTACTAGTACTAATGTTGCTGCTGAG	1089
DB	344	ThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGlu	363
QY	1090	ATCGCGCGCGGATAGCCCTCACCTGTTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTG	1149
DB	364	IleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeu	383
QY	1150	CGCAGAGATTGATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1209
DB	384	ProThrGluLeuSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSer	403
QY	1210	GCCATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAGCAGGATAAG	1269
DB	404	AlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnAspLys	423
QY	1270	GGTATTGCAATCCGCGCATGACCTCGGAGATCTCGTGTGCTTATTCAGGATTAT	1329
DB	424	GlyIleAlaIleProHisAspLeuGlyGluSerArgValValIleGlnAspTyr	443
QY	1330	GATAACCAACATGAACAGATCGCGCGCGCTTCTCCAGCCCATCGCGCGCTTCTCT	1389
DB	444	AspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSer	463
QY	1390	GTCCTTCGAGCTTAATGATGCTTGGCTCTCTCACCGCTGCCAGATGACCGAGTCC	1449
DB	464	ValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyrAspGlnSer	483
QY	1450	ACTTATGGCTCTCGACTCGCGCGCGCTTATGTTTCTGACTCTGCTGCTGCTGCTGCTG	1509
DB	484	ThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnVal	503
QY	1510	GCGACCGCGCGCGCGCGCTTGGCGCTCGCTCGATTGACCAAGGTCACCTTGACGGT	1569
DB	504	AlaThrGlyAlaGlnAlaValAlaAlaArgSerLeuAspTyrThrLysValThrLeuAspGly	523
QY	1570	CGCGCGCTCTCCACCATCGAGTACTCGAACACCTTCTTCTGCTGCTGCTGCTGCTGCTG	1629
DB	524	ArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuProLeuArgGly	543
QY	1630	AAGCTCTCTTCTGGGAGGCGAGCACAACTAAAGCGCGGTACCTTATATATATAACACC	1689
DB	544	LysLeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThr	563
QY	1690	ACTGCTAGCGACCAACTGCTTGTGAGAAATCGCGCGCGCGCGCGCGCGCGCGCGCGCT	1749
DB	564	ThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAlaIleSerThr	583
QY	1750	TACACCATAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1809
DB	584	TyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaValLeuAlaPro	603
QY	1810	CACCTGCTGCTAGCATGCTTGGAGTACCTTGAGTACCTGCTGCTGCTGCTGCTGCTGCTG	1869
DB	604	HisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAlaHisThrPhe	623
QY	1870	GATGATTTCTCCCGAGAGTGGCGCGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGG	1929
DB	624	AspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPheGlnSerThr	643
QY	1930	GTCGCTGAGCTTACGCGCTTACATGAAGTGGGTAAACTCGGAGTTG	1980
DB	644	ValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu	660
RESULT 2			
ID	Q9WQAO	PRELIMINARY; PRT: 660 AA.	
AC	Q9WQAO;		
DT	01-NOV-1999	(TrEMBLrel. 12, Created)	
DT	01-NOV-1999	(TrEMBLrel. 12, Last sequence update)	
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)	
DE	ORF-2.		
OS	Hepatitis E virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage;		
OC	Hepatitis E-like viruses.		
OX	NCBI_TaxID=12461;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=AKL-90;		
RX	MEDLINE=93350000; PubMed=10423137;		
RA	Arankalle V.A., Paranjape S., Emerson S.U., Purcell R.H.,		
RT	Walimbe A.M.;		
RT	"Phylogenetic analysis of hepatitis E virus isolates from India (1976-		
RL	J. Gen. Virol. 80:1691-1700(1999).		
DR	EMBL; AF124407; AAD45493.1; -.		

DR	GO; GO:0005198; F:structural molecule activity; IEA.	
DR	InterPro; IPR004261; SP2.	
DR	InterPro; IPR008975; Viral_cap_coat.	
DR	Pfam; PF03014; SP2; 1.	
SQ	SEQUENCE 660 AA; 70979 MW; DC68116DCD639175 CRC64;	
Alignment Scores:		
Pred. No.:	0	Length: 660
Score:	396.00	Matches: 496
Percent Similarity:	99.80%	Conservative: 0
Best Local Similarity:	99.80%	Mismatches: 1
Query Match:	60.00%	Indels: 0
DB:	12	Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q89468 (1-660)		
QY	313 GCTGGGGCGCGCGCTACCGCGTCCGCGCCATGACACCCCGCGAGTGCCTGAT	372
Db	105 AlaglyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrProValProAsp	124
QY	373 GTGACTCCCGCGCGCCATCTTTCGCGCGCAGTATAACCTATCAACATCTCCCTTACC	432
Db	125 ValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThrSerProLeuThr	144
QY	433 TCTTCGTCGGCCACCGCCTAACCTGCTTCTTATGCGCGCCCTCTTACTGCGCTTTTA	492
Db	145 SerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeu	164
QY	493 CCCCTTCAGACGCGCACCAATACCCATATATATGCGCGCAGGAGCTTCTAATTATGCCAG	552
Db	165 ProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGln	184
QY	553 TACCGGTTGCCCGTGCACAACTCCGTTACCGCGCTGTGTCGCAATGCTGTGCGGGT	612
Db	185 TyrArgValAlaAlaThrIleArgTyrArgProLeuValProAsnAlaValGlyGly	204
QY	613 TAGCCATCTCCATCTCTTCCTGCGCACACACACACACCGCGCTCGCTTGATATG	672
Db	205 TyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAspMet	224
QY	673 AATTCAATAACCTCGACGATGTTTCGTATTTTATGTCAGCGCGCATAGCCTCTGAGCTT	732
Db	225 AsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGluLeu	244
QY	733 GTGATCCCAAGTAGCGCTACACTATCGTAAACCAAGCTGGCGCTCCGTCGAGACCTCT	792
Db	245 ValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThrSer	264
QY	793 GGGGTGCTGAGGAGGCTACCTCTGCTGTTGTTATGCTTGCATACATGGCTCACTC	852
Db	265 GlyValAlaGluGluAlaThrSerGlyLeuValMetLeuGlyIleHisGlySerPro	284
QY	853 GTAATTCCTATATAACACCTATACCGGTGCCCTCGCGCTGTTGGACTTTGCCCTT	912
Db	285 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu	304
QY	913 GAGCTTGAGTTTCGAACCTTACCCCGGTAAACCAATAGCGGGTCTCCGTTATTC	972
Db	305 GluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyrSer	324
QY	973 AGCACTGTCGCCACCGCTTCGTCGGTGCAGCGGACTGCCGAGCTCACCAACACG	1032
Db	325 SerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGluLeuThrThrThr	344
QY	1033 GCTGCTACCCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGTCGTCGTCGATC	1092
Db	345 AlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGluIle	364
QY	1093 GGCGCGGATAGCCCTCACCTGTTCACCTTCTGCTACCTCTGCTGTCGTCGCGCTCGC	1152
Db	365 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuPro	384
QY	1153 ACAGATTGATTTCGTCGGCTGGTGCAGCTGTCTACTCCGTCGCTGCTCTCAGCC	1212
Db	385 ThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAla	404
QY	1213 AATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGACAAGTCTCAGCAGGATAGGCT	1272
Db	405 AsnGlyGluProThrValIysLeuTyrThrSerValGluAsnAlaGlnGlnAspIysGly	424
QY	1273 ATTGCAATCCCGCATGACATTCACCTCGGAGAAATCTCGTGTGTTATTTCAGGATTATGAT	1332
Db	425 IleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIleGlnAspTyrAsp	444
QY	1333 AACCAACATGAACAAGATCGCGCGCCTTCTCCAGCCCCATCGCGCCCTTCTCTCTC	1392
Db	445 AsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerVal	464
QY	1393 CTTGAGACTAATGATGTGTTGGCTCTCTCTCACCGCTCCGAGTATGACCACTCCACT	1452
Db	465 LeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThr	484
QY	1453 TATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTGGTTAATGTCG	1512
Db	485 TyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAla	504
QY	1513 ACCGGCGCGCAGCCCTTCCCGCTCGCTCGATTGGAACCAAGTGCACCTTGACGTCGC	1572
Db	505 ThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArg	524
QY	1573 CCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCTCGCGCTCCGCGTAA	1632
Db	525 ProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuProLeuArgGlyLys	544
QY	1633 CTCTCTTCTGGAGGCGGACCAACTAAAGCCGGGTACCTTTAATATAACACCACT	1692
Db	545 LeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThr	564
QY	1693 GCTAGCGACCAACTGTTGTCGAGATGCGCGCGGACCGGGTCGCTATTTCCACTTAC	1752
Db	565 AlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyr	584
QY	1753 ACCACTAGCTGGTCTGCTCGCTCCGCTCCATTTCTGCGGTTCGCTTTTA 1803	
Db	585 ThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaValLeu 601	
RESULT 4		
Q9W7W9	PRELIMINARY; PRT; 660 AA.	
ID	Q9W7W9	
AC	Q9W7W9; 01-NOV-1999 (TrEMBLrel. 12, Created)	
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	Structural protein 2 (Structural viral protein).	
OS	Hepatitis E virus.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage;	
OC	Hepatitis E-like viruses.	
OX	NCBI_TaxID=12461;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=HEV-Morocco;	
RX	MEDLINE=99107414; PubMed=9892396;	
RA	Meng J., Cong M., Dai X., Pillot J., Purdy M.A., Fields H.A.,	
RA	Khudyakov Y.E.;	
RT	"Primary structure of open reading frame 2 and 3 of the hepatitis E	
RT	virus isolated from Morocco.";	
RL	J. Med. Virol. 57:126-133(1999).	
RL	[2]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=Morocco;	
RC	STRAIN=Morocco;	
RA	Chen G., Meng J.;	
RT	"Identification of the 5' Capped and 3' Complete Terminal Sequence of	
RT	the Hepatitis E virus Isolated From Morocco.";	
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF065061; A010627.1; -.	
DR	EMBL; AY230202; A072992.1; -.	

DR GO: 0005198; F: structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70943 MW; 9719D30CCD932950 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 660
Score: 362.00 Matches: 562
Percent Similarity: 99.65% Conservatives: 0
Best Local Similarity: 99.65% Mismatches: 2
Query Match: 54.85% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q9W7M9 (1-660)

```
QY 289 GCCTCAGTCGTAGACCTACCAAGCTGGGGCGCGCGCTAACCGCGTGCCTCGGCC 348
DB 97 AlaSerArgArgProThrThrAlaGlyAlaAlaProLeuThrAlaValaProAla 116
QY 349 CATGACACCCGCCAGTGCCTGATGCGACTCCCGCGCGCCATCTTGGCGCGGAGTAT 408
DB 117 HisAspThrProProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyr 136
QY 409 AACCTATCAACATCTCCCTTACCTCTTCCGTGGCCACCGCACTAACCTGGTCTTTAT 468
DB 137 AsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyr 156
QY 469 GCCGCCCTCTTAGTCGGCTTTTACCCCTTAGGACGGCACCACCAATACCATATATGGCC 528
DB 157 AlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAla 176
QY 529 ACGAAGCTTCTTAATTATGCCAGTACCGGGTGGCGGCCACCAATCGTTACCGCCG 588
DB 177 ThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgPro 196
QY 589 CTGGTCCCCAATGCTGTGGCGGTGTACGCCATCTCCATCTCTGCGCACAGACCCACC 648
DB 197 LeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThr 216
QY 649 ACCACCGGACGTCGTTGATATGAAATCAATCACTCGACGGATGTCGATTTTATGTC 708
DB 217 ThrThrProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuVal 236
QY 709 CAGCCCGGCATAGCTCTGAGCTGTGATCCCAAGTACGCGCTACCTATCGTAACCAA 768
DB 237 GlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGln 256
QY 769 GGCTGGCGCTCCGTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACCTCTGTTGTT 828
DB 257 GlyTyrArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuVal 276
QY 829 ATGCTTTGCATACATGGCTCACTCGTAATTCCTATACCTATACCACTATACCGGTGCC 888
DB 277 MetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAla 296
QY 889 CTCGGGCTGGTGGACTTTGGCCCTTCAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCC 948
DB 297 LeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsnLeuThrProGlyAsnThr 316
QY 949 AATACGGCGGTCTCCCGTTATTCAGCACTGTCTGCCACCGCTTCTCGTCGGTGGCAG 1008
DB 317 AsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAsp 336
QY 1009 GGGACTGGCGAGCTCACCAACGAGTGTACCCGCTTTATGAAGACCTCTATTTTACT 1068
DB 337 GlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheSer 356
QY 1069 AGTACTAATGGTGTGGTGAAGTACGGCGCGGATAGCCCTACCTGTTCAACCTTGTCT 1128
DB 357 SerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAla 376
QY 1129 GACACTCTGCTTGGCGGCGCTGCGGACAGAAATTGATTTGCTCGGCTGGTGGCGAGCTGTC 1188
```

```
DB 377 AspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPhe 396
QY 1189 TACTCCCGTCCGTTGCTCTACGCCAATGCGAGCGGACTGTTAAGTTGATACATCTGA 1248
DB 397 TyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal 416
QY 1249 GAGATGCTCAGCAGGATAAGGTATTGCAATCCCGCATGACATTACCTCGGAGATCT 1308
DB 417 GluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSer 436
QY 1309 CGTGTGGTTATTACAGGATTATGATTAACCAACATGACAGATCGGCCGAGCTTCTCCA 1368
DB 437 ArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 456
QY 1369 GCCCATCGCGCCCTTCTCTCTCTCGAGCTAATGATGTGCTTGGCTCTCTCTCAC 1428
DB 457 AlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThr 476
QY 1429 GCTGCGGAGTAGACCGAGTCCACTTATGCTCTTCGACTGGCCCGCAGTTTATCTTCTGAC 1488
DB 477 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 496
QY 1489 TCTGTGACCTTGGTTAATGTTTTCGACCGCGCGCAGCGCTTCCCGGTCGCTCGATTGG 1548
DB 497 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyr 516
QY 1549 ACCAAGGTCACTTGACGTCGCGCTCTCTCCACATCCAGCAGTACTCGAAGACCTTC 1608
DB 517 ThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhe 536
QY 1609 TTTGTCCTGCGCTCCGCGGTAAGCTCTCTTCTCGGAGGAGGACACACTAAAGCCGG 1668
DB 537 PheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGly 556
QY 1669 TACCTCTTAATTAACACCACTGCTAGCAGCACTGCTTGTGAGAAATGCCCGCGG 1728
DB 557 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGly 576
QY 1729 CACCGGTCGCTATTTCCACTTACACCTAGCTGGGTGGTGGTCCGCTCTCCATTCT 1788
DB 577 HisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSer 596
QY 1789 GCGGTTGCGTTTAGCCCCCCTCTGCTAGCATTTGCTGAGGATACCTTGACTAC 1848
DB 597 AlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyr 616
QY 1849 CCGTCCCGCGCCCATACATTTTGATGATTTCTGCCAGAGTGGCGCCCTTGGCTTCAG 1908
DB 617 ProAlaArgAlaHisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGln 636
QY 1909 GGCTCGCTTTCCAGTCTACTGTCGCTGAGCTTACAGCCCTTAAGATGAAGTGGGTAAA 1968
DB 637 GlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLys 656
QY 1969 ACTCGGAGGTG 1980
DB 657 ThrArgGluLeu 660
RESULT 5
Q9WJZ5
ID Q9WJZ5 PRELIMINARY; PRT; 344 AA.
AC Q9WJZ5;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Structural protein 2 (fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN=2B Abbottabad;
RA Van Cuyck-Candre H., Clements N.J., Zhang H.Y., Caudill J.C.,
RA Cohen S.G., Coursaget P., Buisson Y., Warren R.L., Longer C.F.;
RT "Partial Sequence of HEV Isolates from North Africa and Pakistan;
RT Comparison with Known HEV Sequences.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40044; AAD09445.1; -;
DR GO; GO:0005198; F.structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR Pfam; PF03014; SP2; 1.
FT NON TER 1 1
SQ SEQUENCE 344 AA; 37146 MW; D7F3D2D9BB44098B CRC64;

Alignment Scores:
Pred. No.: 0 Length: 344
Score: 344.00 Matches: 344
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.12% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q9WJZ5 (1-344)

QY 949 AATACCGGGCTCTCCGGTTATTTCAGCACTGCTCGCCACCGCTTCGTCGGGTGGGAC 1008
DB 1 AsnThrArgValSerArgTyrSerThrAlaArgHisArgLeuArgGlyAlaAsp 20

QY 1009 GGGACTCGCGAGTCCACCACCGGTGCTACCGCTTATGAGGACCTCTATTACT 1068
DB 21 GlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThr 40

QY 1069 AGTACTAATGGTCTCGGTGAGTCGGCGGGATAGCCCTCAACCTGTTCAACCTTGCT 1128
DB 41 SerThrAsnGlyValGlyGluLeuGlyArgGlyLeuAlaLeuThrLeuPheAsnLeuAla 60

QY 1129 GACACTCTGCTCGCGGCTCCGCGACAGAAATGATTTCGTCGGCTGGTCCAGCTGTTT 1188
DB 61 AspThrLeuLeuGlyGlyLeuProThrGluLeuLeuSerSerAlaGlyGlyGlnLeuPhe 80

QY 1189 TACTCCGCTCCGTTGCTCAGCAATGGCGAGCGAGCTTAAAGTTGTATACATCTGTA 1248
DB 81 TyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal 100

QY 1249 GACAATGCTCAGCAGGATAGGTTATGCAATCCGCAATGACATGACCTCGCAGAACT 1308
DB 101 GluAsnAlaGlnGlnAspLysGlyLeuAlaIleProHisAspLeuGlyGluSer 120

QY 1309 CGTGTGGTTATTCAGGATTATGATACCAACATCAAGATCGCGCGCTTCTCCA 1368
DB 121 ArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 140

QY 1369 GCCCATCGCGCTTCTCTGCTCTCGAGCTAAATGATGCTTGGCTCTCTCTCACC 1428
DB 141 AlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThr 160

QY 1429 GCTGCGCAGATACACAGTCCACTTATGCTCTTCGACTCGCCAGTTTATGTTCTGAC 1488
DB 161 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 180

QY 1489 TCTGTGACCTTGGTTAATGTTGCGACCGCGCGCGCCGTTGCCGGTTCGCTCGATTGG 1548
DB 181 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyr 200

QY 1549 ACCAAGGTACACTTGAGCGTCCGCCCTCTCCACCATCCAGCACTGCTCGAGACCTTC 1608
DB 201 ThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhe 220

QY 1609 TTTGTCTCGCTCCCGGTAACTCTCTTTCTGGGAGGAGCAGCAACAATAAGCGCGG 1668
DB 221 PheValLeuProLeuArgGlyLysLeuSerPheThrProGluAlaGlyThrThrLysAlaGly 240

QY 1669 TACCTTATATTAACACCACTGTGCTAGCGACCAACTGCTTGTGAGAAATGCGCGCGG 1728

DB 241 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGly 260
QY 1729 CACCGGTGCTATTTCACCTTACACTAGCTAGCTGGGCTGGTCCGCTCTCCATTCT 1788
DB 261 HisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSer 280
QY 1789 GCGGTTGCGCTTTAGCCCCCACTCTGCGCTAGCATTCCTTGAGGATACCTTGGACTAC 1848
DB 281 AlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuGluAspThrLeuAspTyr 300
QY 1849 CTGCGCGCGCCCATACTTTTGTATGATTTCTGCGCCAGAGTGGCGCCCTTGGCCTTCAG 1908
DB 301 ProAlaArgAlaHisThrPheAspPheCysProGluCysArgProLeuGlyLeuGln 320
QY 1909 GCGTGGCTTCCAGCTCTACTGCTGAGCTTCAGCGCTTAAGATGAAGTGGGTAAA 1968
DB 321 GlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLys 340
QY 1969 ACTCGGGAGTTG 1980
DB 341 ThrArgGluLeu 344

RESULT 6
Q81871 PRELIMINARY; PRT; 660 AA.
AC Q81871; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF 2 precursor.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=92024067; PubMed=1926770;
RX Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Kikihisa T., Mizuno K., Win K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=92335008; PubMed=1630924;
RA Aye T.T., Uchida T., Ma X.Z., Iida F., Shikata T., Zhuang H.,
RA Win K.M.;
RT "Complete nucleotide sequence of a hepatitis E virus isolated from the
RT Xinjiang epidemic (1986-1988) of China.";
RL Nucleic Acids Res. 20:3512-3512(1992).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
[5]
RN SEQUENCE FROM N.A.
RX MEDLINE=92271462; PubMed=1589964;
RA Fry K.E., Tam A.W., Smith M.W., Kim J.P., Luk K.-C., Young L.M.,
RA Piatak M., Feldman R.A., Yun K.Y., Purdy M.A., McCaustland K.A.,
RA Bradley D.W., Reyes G.R.;
RT "Hepatitis E virus (HEV): Strain variation in the nonstructural gene
RT region encoding consensus motifs for an RNA-dependent RNA polymerase
RT and an ATP/GTP binding site.";

RL Virus Genes 6:173-185 (1992).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93348763; PubMed=8346669;
 RA Bi S-L., Purdy M.A., McCaustland K.A., Margolis H.S., Bradley D.W.;
 RT "The sequence of hepatitis E virus isolated directly from a single
 source during an outbreak in China."
 RL Virus Res. 28:233-247 (1993).
 DR EMBL; L08816; AAA03191.1; -;
 GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR004261; SP2.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF03014; SP2; 1.
 KW Signal.
 FT SIGNAL
 SQ SEQUENCE 660 AA; 70979 MW; B58F23955FDD6614 CRC64;
 Alignment Scores:
 Pred. No.: 3,64e-274 Length: 660
 Score: 284.00 Matches: 284
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 43.03% Indels: 0
 DB: 12 Gaps: 0
 US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q81871 (1-660)
 QY 733 GTGATCCCAAGTGGAGCGCTACATATCGTAACCAAGCGTGGCGCTCGGTGAGACCTCT 792
 Db 245 ValileProserGluArgLeuHisTyArgAsnGlnGlyTrpArgSerValGluThrSer 264
 QY 793 GGGTGTCTGAGGAGGAGCTACTCTGCTGTCTTGTATGCTTTGCATACATGCTCACTC 852
 Db 265 GlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerLeu 284
 QY 853 GTAAATTCCTACTATACACCTATACCGGTGCCCTCGGCTGTGGACTTGGCCCTT 912
 Db 285 ValAsnSerTyThrAsnThrProTyThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 304
 QY 913 GAGCTTCAGTTCGCAACCTTACCCCGGTAAACCAATACCGGGTCTCCCGTTATTCC 972
 Db 305 GluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTySer 324
 QY 973 AGCACTCTCGCACCGCCCTTGGTCGGGTGCGGACGGAGTCCGAGCTCACCACCAACG 1032
 Db 325 SerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGluLeuThrThr 344
 QY 1033 GCTGCTACCGCTTATGAGGACCTCTATTTACTAGTACTAATGTTGTCGGTGAGATC 1092
 Db 345 AlaAlaThrArgPheMetLysAspLeuTyPheThrSerThrAsnGlyValGlyGluIle 364
 QY 1093 GGGCGGGGATAGCCCTCACCCCTGTCAACCTTGCTGACACTCTGCTTGGCGGCTGCCG 1152
 Db 365 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuPro 384
 QY 1153 ACAGAAATGATTTCGTGGTGGTGGCCAGCTGTCTTACTCCGTCGCCGTGCTCAGCC 1212
 Db 385 ThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTySerArgProValValSerAla 404
 QY 1213 AATGGCGAGCGGCTGTAGTTGTATACATCTGTAGAGATGCTCAGCAGGATAGGGT 1272
 Db 405 AsnGlyGluProThrValLysLeuTyThrSerValGluAsnAlaGlnGlnAspLysGly 424
 QY 1273 ATTGCAATCCCGCATGATCAATTGACTCCGAGAACTCTCGTGTGGTTATTACAGGATTATGAT 1332
 Db 425 IleAlaIleProHisAspIleAspLeuGlyGlnSerArgValValIleGlnAspTyAsp 444
 QY 1333 AACCAACATGAACAAGATCGCGCCGACGCCCTTCTCCAGCCCCATCGCGCCCTTCTCTGTC 1392
 Db 445 AsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerVal 464
 QY 1393 CTTGCGAGCTAATGATGTGCTTGGCTCTCTCTCACCGCTCCGAGTATGACAGTCCACT 1452

```
Db      365 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuPro 384
QY      1153 ACAGAAATTGATTCGTGGCTGGTGGCCAGCTGTTCTACTCCGTCGGTTCCTCAGCC 1212
Db      385 ThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAla 404
QY      1213 AATGGGAGCCGACTCTTAAGTTGTATACATCTGTAGAAATGCTCAGCAGGATAAGGGT 1272
Db      405 AsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnAspLysGly 424
QY      1273 ATTGCAATCCCGCATGACATTCACCTCGGAGAAATCTCGTGTGTTATTACAGGATTATGAT 1332
Db      425 IleAlaIleProHisaspIleaspLeuGlyGluSerArgValValIleGlnAspTyrAsp 444
QY      1333 AACCAACATGAACAAGATCGCCGAGCCCTTCTCCAGCCCATCGGCCCTTCTCTGTC 1392
Db      445 AsnGlnHisgluGlnAspArgProThrProSerProAlaProSerArgProPheSerVal 464
QY      1393 CTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTCCGAGTATGACACAGTCCACT 1452
Db      465 LeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGlnTyrAspGlnSerThr 484
QY      1453 TATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACTTGTGCTTGTGTTAATGTGG 1512
Db      485 TyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAla 504
QY      1513 ACCGGCGCGAGCGCGTTCGGCTCGCTCTCTCACCGCTCCGAGTATGACACAGTCCACT 1572
Db      505 ThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArg 524
QY      1573 CCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTCTCGCGCTCCGCGGTAAG 1632
Db      525 ProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuProLeuArgGlyLys 544
QY      1633 CTCTCTTTCTGGAGGAGGAGCACAACTAAAGCCGGGTACCTTTATATTATTAACACCACT 1692
Db      545 LeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThr 564
QY      1693 GCTAGC 1698
Db      565 AlaSer 566
```

RESULT 8

```
Q8UT32 PRELIMINARY; PRT; 659 AA.
ID Q8UT32 AC Q8UT32;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=22072932; PubMed=12076829;
RA Jameel S., Zafrullah M., Chawla Y.K., Dilawari J.B.;
RT "Reevaluation of a North India isolate of hepatitis E virus based on
RT the full-length genomic sequence obtained following long RT-PCR.";
RL Virus Res. 86:53-58 (2002).
RN [2]
RN SEQUENCE FROM N.A.
RA Jameel S., Zafrullah M., Chawla Y.K., Dilawari J.B.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF459438; AAM66330.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 659 AA; 70842 MW; ABCC177EBE6012C8 CRC64;
```

```
Alignment Scores: 3.58e-265 Length: 659
Pred. No.: 275.00 Matches: 375
Score: 99.73% Conservative: 0
Percent Similarity: 99.73% Mismatches: 1
Best Local Similarity: 41.67% Indels: 0
Query Match: 12 Gaps: 0
DB:

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q8UT32 (1-659)
QY 853 GTAAATCTTACTAATAATACACCTTACGGTGCCTCGGCTGTGTGGACTTTGCCCTT 912
Db 284 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 303
QY 913 GAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGGTCTCCCTTTATTC 972
Db 304 GluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyrSer 323
QY 973 AGCACTGCTCGCCACCGCTTCGTCCGCTGCGGACGGGACTCCCGAGCTCACCAACG 1032
Db 324 SerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGluLeuThrThr 343
QY 1033 GCTGCTACCGCTTTTATGAGGACCTCTATTTTACTAGTACTAATGTTGCGTGAGATC 1092
Db 344 AlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGluIle 363
QY 1093 GGCGCGGGATAGCCCTCACCCCTGTTCAACCTTGTCTGACACTCTGTGTCGGCCCTGCCG 1152
Db 364 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuPro 383
QY 1153 ACAGAAATTGATTCGTGGCTGGTGGCCAGCTTCTACTCCGTCGGTTCCTCAGCC 1212
Db 384 ThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAla 403
QY 1213 AATGGGAGCCGACTCTTAAGTTGTATACATCTGTAGAAATGCTCAGCAGGATAAGGGT 1272
Db 404 AsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnAspLysGly 423
QY 1273 ATTGCAATCCCGCATGACATTCGCTCGGAGAAATCTCGTGTGTTATTACAGGATTATGAT 1332
Db 424 IleAlaIleProHisaspIleaspLeuGlyGluSerArgValValIleGlnAspTyrAsp 443
QY 1333 AACCAACATGAACAAGATCGCGGAGCCCTTCTCCAGCCCATCGGCCCTTCTCTGTC 1392
Db 444 AsnGlnHisgluGlnAspArgProThrProSerProAlaProSerArgProPheSerVal 463
QY 1393 CTTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTCCGAGTATGACACAGTCCACT 1452
Db 464 LeuArgAlaAsnAspValLeuTrpLeuSerLeuThrGlyAlaGlnTyrAspGlnSerThr 483
QY 1453 TATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACTTGTGTTAATGTGGTGG 1512
Db 484 TyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAla 503
QY 1513 ACCGGCGCGAGCGCGTTCGGCTCGCTCGATTGACCAAGTTCACACTTACAGGTCGC 1572
Db 504 ThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArg 523
QY 1573 CCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTCGCGCTCCGCGGTAAG 1632
Db 524 ProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuProLeuArgGlyLys 543
QY 1633 CTCTCTTTCTGGAGGAGGAGCACAACTAAAGCCGGGTACCTTTATATTATTAACACCACT 1692
Db 544 LeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThr 563
QY 1693 GCTAGCACCACACTGCTTGTTCGAGAAATGCCCGCGGACCGGGTCGCTATTTCACACTAC 1752
Db 564 AlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyr 583
QY 1753 ACCACTAGCTGGGTGCTGCTCCCGTCTCATTTCTCGGGTTCGGTTTAGCCCCCCCAC 1812
Db 584 ThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaValLeuAlaProHis 603
```

QY 1813 TCTGGCTAGCATTCCTTGTAGGATACCTTGGACTACCTGCGCGCCCATCTTTTGTAT 1872
 Db 604 SerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAlaHisThrPheAsp 623
 QY 1873 GATTTCCTCCAGAGTGGCGCCCTTGGCTTCCAGGCTGGCTTCCAGTCTACTGTC 1932
 Db 624 AspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrVal 643
 QY 1933 GCTGAGCTTCAGCGCTTAAGATGAAGTGGTAAACTCGGAGTTG 1980
 Db 644 AlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 659
 RESULT 9
 Q69411 ID Q69411 PRELIMINARY; PRT; 660 AA.
 AC Q69411;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Structural protein.
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=hev037;
 RA Donati M.C., Fagan E.A., Harrison T.J.;
 RT "Sequence analysis of full length HEV clones derived directly from
 human liver in fulminant hepatitis E.";
 RL (In) Rizzetto M., Purcell R.H., Gerin J.L., Verme G (eds.);
 RL VIRAL HEPATITIS AND LIVER DISEASE, pp.313-316,
 RL Edizioni Minerva Medica, Torino (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=hev037;
 RA Harrison T.J.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X98292; CAA66937.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR004261; SP2.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF03014; SP2; 1.
 SQ SEQUENCE 660 AA; 71024 MW; F37E3FEFF7A4EAD1 CRC64;
 Alignment Scores:
 Pred. No.: 3, 52e-254 Length: 660
 Score: 264.00 Matches: 264
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 40.00% Indels: 0
 DB: 12 Gaps: 0
 US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q69411 (1-660)
 QY 919 GAGTTTCGCAACCTTACCCCGGTACACCAATACGCGGTCTCCGCTATTTCAGCACT 978
 Db 307 GluPheArgAsnLeuThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThr 326
 QY 979 GCTCGCCACGCTTCGTGCGGTGGGAGCGGAGTCCGAGTCCACCAACGCGTGTCT 1038
 Db 327 AlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAla 346
 QY 1039 ACCCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGGTCTCGTGAGATCGCGCGC 1098
 Db 347 ThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGluLeuGlyArg 366
 QY 1099 GGGATAGCCCTCACCTGTTTCAACCTTGTGACACTCTGCTGGCGGCTGGCGCAGAA 1158
 Db 367 GlyValAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGlu 386
 QY 1159 TTGATTTCTGCGCTGGTGGCCAGCTGTTTCTACTCCCGTCCCGTGTCTCAGCCAATGGC 1218

Db 387 LeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGly 406
 QY 1219 GAGCGCACTGTTAGTTGTATACATCTGTAGAAATGCTCAGCAGATAAGGTTATTGCA 1278
 Db 407 GluProThrValLysLeuTyrThrSerValGluAsnAlaGlnAspLysGlyIleAla 426
 QY 1279 ATCCCGCATGACATTGACCTCGGAGAACTCTCGTGGTGTATTTCAGGATTATGATACCAA 1338
 Db 427 IleProHisAspIleAspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGln 446
 QY 1339 CATGAACAAGATCGCGCGAGCCCTTCTCCAGCCCCCATCGGCCCTTCTCTGCTCTCGA 1398
 Db 447 HisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeuArg 466
 QY 1399 GCTAATGATGCTTTGGCTCTCTCTCACCCTGCCGAGTATGACCATGCTTATGGC 1458
 Db 467 AlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGly 486
 QY 1459 TCTTCGACTGCGCCAGTTATGTTCTGACTCTGCTGACCTTGGTAAATGTTGCGACCGGC 1518
 Db 487 SerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAlaThrGly 506
 QY 1519 GCGCAGCGCTTGCCCGGTGCTCGATTGGACCAAGGTACACTTACGCTGCGGCCCTC 1578
 Db 507 AlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeu 526
 QY 1579 TCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTCGCGCTCCGCGGTAAGCTCTCT 1638
 Db 527 SerThrIleGlnTyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSer 546
 QY 1639 TTCTGGGAGCAGCACAACCTAAGCCGGTACCTTATATATATACACCATCTGCTAGC 1698
 Db 547 PheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSer 566
 QY 1699 GACCACTGCTT 1710
 Db 567 AspGlnLeuLeu 570
 RESULT 10
 Q69419 ID Q69419 PRELIMINARY; PRT; 660 AA.
 AC Q69419;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE ORF1, ORF2 & ORF3.
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=12461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA von Brunn A., Seebach J., Thyagarajan S.P., Mohanavalli B., Menon T.,
 RA Froesner G.;
 RT "PCR amplification, cloning and sequence determination of a hepatitis
 E virus isolate from Madras, India.";
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X99441; CAA67804.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR004261; SP2.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF03014; SP2; 1.
 SQ SEQUENCE 660 AA; 70960 MW; DAC3DF95F91689F8 CRC64;
 Alignment Scores:
 Pred. No.: 3, 49e-249 Length: 660
 Score: 259.00 Matches: 559
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 3
 Query Match: 39.24% Indels: 0
 DB: 12 Gaps: 0

QY	1375	TCGCGCCCTTCTCTCTCTCTCGAGCTAATGATGCTTGGCTCTCTCTCACCGCTGCC	1433
Db	459	SerArgProPheSerAlaLeuArgAlaAsnAspValLeuIrpLeuSerLeuThrAlaAla	478
QY	1435	GAGTATGACACAGTCCACTTATGGCTCTTCGACTGGCGCCAGCTTTATGTTTCTGACTCTGTG	1494
Db	479	GluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerVal	498
QY	1495	ACCTTGGTTAAATGTCGGACCGCGCGGAGGCGGTTCGCCGTGCTCGATTCGAGCAACAG	1554
Db	499	ThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaAArgSerLeuAspTrpThrLys	518
QY	1555	GTCCACTTTCACCGGTGCGCCCTCTCCACCATCCAGCAGTACGAGACCTTCTTTGTC	1614
Db	519	ValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheVal	538
QY	1615	CTCCGCTCCGCGGTAACTCTCTTCTGGGAGGCGAGGACCAACTAAAGCGGGTACCCT	1674
Db	539	LeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrPro	558
QY	1675	TATAATTATTAACACCACTGCTACGAGCAACTGCTGTGCGAGAAATCGCGCGGCACCGG	1734
Db	559	TyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArg	578
QY	1735	GTCGCTATTTCCTACTACACCATAGCTCGGTGGTGGTCCCGTCTCCATTTCTGCGGTT	1794
Db	579	ValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaVal	598
QY	1795	GCCGTTTTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGGACTACCCCTGCC	1854
Db	599	AlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAla	618
QY	1855	CGCGCCCATACTTTTGATGATTTCTGCCAGAGTGGCGCCCTTGGCTTCTAGGGTCTGC	1914
Db	619	ArgAlaHisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCys	638
QY	1915	GCITTCAGCTTACTGTCGTGAGCTTCAGGCTTAAGACTGAAGTGGGTGTAACCTCGG	1974
Db	639	AlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArg	658
QY	1975	GAGTTG 1980	
Db	659	GluLeu 660	
RESULT 11			
ID	O91855	PRELIMINARY; PRT; 660 AA.	
AC	O91855;		
DT	01-NOV-1998 (TrEMBLrel. 08, Created)		
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Structural protein.		
OS	Hepatitis E virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage;		
OC	Hepatitis E-like viruses.		
OX	NCBI_taxID=12461;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=93-Egypt;		
RE	MEDLINE=99105430; PubMed=9890424;		
RA	Tsarev S.A., Binn L.N., Gomatos P.J., Arthur R.R., Monier M.K.,		
RA	van Cuyck-Gandre H., Longer C.F., Innis B.L.;		
RT	"Phylogenetic analysis of hepatitis E virus isolates from Egypt.;"		
RL	J. Med. Virol. 57:68-74(1999).		
DR	EMBL; AF051351; AAC35761.1; -		
DR	GO; GO:0005138; F:structural molecule activity; IEA.		
DR	InterPro; IPR004261; SP2.		
DR	InterPro; IPR008975; Viral_cap_coat.		
DR	Pfam; PF03014; SP2; 1.		
SO	SEQUENCE 660 AA; 71040 MW; 044FF5EA7C492791 CRC64;		

Alignment Scores:

Pred. No.: 3,44e-241 Length: 660
 Score: 251.00 Matches: 551
 Percent Similarity: 99.46% Conservative: 0
 Best Local Similarity: 99.46% Mismatches: 3
 Query Match: 38.03% Indels: 0
 DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x 091855 (1-660)

QY 31 CTCAGTGTTCCTATAGTCTGCCCGCCACCGCCCGGTACGCTGTGGCCGCGTGT 90
 Db 11 LeuMetPheLeuProMetLeuProAlaProProGlyGlnProSerGlyArgArg 30
 QY 91 GGGCGGCGCAGCGGGTCCGGCGGTCTCTGGGGTACCGGGTGTGATCTCAGCCC 150
 Db 31 GlyArgSerGlyGlySerGlyGlyGlyPheTrpGlyAspValAspSerGlnPro 50
 QY 151 TTCGCAATCCCTATATTCATCCCAACCAACCCCTTCGCCGCCCGATGTCCAGCCG 210
 Db 51 PheAlaIleProTyrlleHisProThrAsnProPheAlaProAspValThrAlaAla 70
 QY 211 GGGGTGACCTCGTGTTCGCCAACCGCCGACCACTCGGCTCGCTTGGGTGACCA 270
 Db 71 GlyAlaGlyProArgValArgGlnProAlaArgProLeuGlySerAlaTrpArgAspGln 90
 QY 271 GCCAGAGCCCGCGGTTCCTCAGTCTGAGACCTACACAGCTGGGGCGCGCGCTA 330
 Db 91 AlaGlnArgProAlaValAlaSerArgArgProThrThrAlaGlyAlaAlaProLeu 110
 QY 331 ACCGGGTGCTCCGCCGCCATGACACCCCGCCAGTGCCTGTGACTCCCGCGCGCC 390
 Db 111 ThrAlaValAlaProAlaHisAspThrProProValProAspValAspSerArgGlyAla 130
 QY 391 ATCTTGGCGCGCAGTATACTATCAACATCTCCCTTACCTCTTCCTGGTGGCCACCG 450
 Db 131 IleLeuArgArgGlnTyAsnLeuSerThrSerProLeuThrSerSerValAlaThrGly 150
 QY 451 ACTACCTGGTCTCTTATGTCGCCCGCTCTTGTAGTCGCTTTTACCTTCAGGACGGCAC 510
 Db 151 ThrAsnLeuValLeuTyAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThr 170
 QY 511 AATACCCATATATGCGCCAGGAAGTCTCTAATTATGCCAGTACCGGTGTCCCGTGC 570
 Db 171 AsnThrHisIleMetAlaThrGluAlaSerAsnTyAlaGlnTyArgValValArgAla 190
 QY 571 ACAATCCGTTACGCGCGTGTGCCAATGCTGTGCGGGTGTAGCCATCTCCATCTCA 630
 Db 191 ThrIleArgTyArgProLeuValProAsnAlaValGlyGlyTyAlaIleSerIleSer 210
 QY 631 TTCTGCGCACAGACACACCCCGCGCTCGTGTGATGATGAATTCATAACCTCGACG 690
 Db 211 PheTrpProGlnThrThrThrThrProThrSerValAspMetAsnSerIleThrSerThr 230
 QY 691 GATGTTCGTATTTAGTCCAGCCCGGCATAGCCTCTGAGCTTGTGATCCCAAGTGAAGC 750
 Db 231 AspValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArg 250
 QY 751 CTACACTATGTAAACAGGCTGGCGCTCCGTCGAGACCTCTGGGTGGCTGAGGAGAG 810
 Db 251 LeuHisTyArgAsnGlnGlyTrpArgSerValGluThrSerGlyValAlaGluGluGlu 270
 QY 811 GCTACCTCTGGTCTGTATGCTTTCATACATGCGCTCACTCGTAATTCCTATATAAT 870
 Db 271 AlaThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyThrAsn 290
 QY 871 ACACCCATATACCGGTGCTCGCGCTGTGGACTTTGGCCCTTCAGCTTTCAGTTTCGCAAC 930
 Db 291 ThrProTyThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsn 310
 QY 931 CTTACCCCGGTAAACCAATACCGGGTCTCCCGTATTTCAGCACTGCTCCGCCACCGC 990
 Db 311 LeuThrProGlyAsnThrAsnThrArgValSerArgTySerSerThrAlaArgHisArg 330

QY 991 CTTCGTCCGCGTGGCGACGGAGCTCCGAGCTCACACCGGCTGTACCGCTTTATG 1050
 Db 331 LeuArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMet 350
 QY 1051 AAGGACCTCTATTTTACTAGTACTAATGCTGCGTGAGATCGCGGGATAGCCCTC 1110
 Db 351 LysAspLeuTyPheSerSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeu 370
 QY 1111 ACCCTGTTCAACCTTGCTGACACTCTGCTTCGCGCCTCGCGACAGAAATGATTTTCGTG 1170
 Db 371 ThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSer 390
 QY 1171 GCTGTGCGCAGCTGTTCTTACTCCCGTCCCGTGTCTCAGCAATGGCGAGCCGACTGT 1230
 Db 391 AlaGlyGlyGlnLeuPheTySerArgProValValSerAlaAsnGlyGluProThrVal 410
 QY 1231 AAGTTGTATACATCTGTAGAGATGCTCAGCAGGATAGGATTCGATCCCGCATGAC 1290
 Db 411 LysLeuTyThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAsp 430
 QY 1291 ATTGACCTCGAGAAATCTCGTGTGTTTATTCAGGATTTATGATTAACCAACATGAAACAAGAT 1350
 Db 431 IleAspLeuGlyGluSerArgValValIleGlnAspTyAspAsnGlnHisGluGlnAsp 450
 QY 1351 CGCCGCGACCTTCTCCAGCCCGCATCGCGCCCTTTCTCTGCTCTTCGAGCTAATGATGTG 1410
 Db 451 ArgProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspVal 470
 QY 1411 CTTTGGCTCTCTCACCGCTGCCGAGTATGACAGTCACTTATGCTTTCGCTTCGAGTGGC 1470
 Db 471 LeuTrpLeuSerLeuThrAlaAlaGluTyArgGlnSerThrTyArgTyArgSerThrGly 490
 QY 1471 CCAGTTTATGTTCTGACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGCAGCCGCTT 1530
 Db 491 ProValTyValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaVal 510
 QY 1531 GCCCGGTGCTCGATTGGACCAAGGTCACTTGACGTCGCGCCCTCTCCACCATCCAG 1590
 Db 511 AlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGln 530
 QY 1591 CAGTACTCGAAGACCTTCTTGTCTGCGCTCCCGCTCCCGGTAAAGTCTCTTTCTGGGAGGCA 1650
 Db 531 GlnTySerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAla 550
 QY 1651 GSCACAACTAAAGCGGGTACCCTTATATTAATAACACCACT 1692
 Db 551 GlyThrThrLysAlaGlyTyProTyArgAsnTyAsnThrThr 564

RESULT 12
 QYWL0 PRELIMINARY; PRT; 605 AA.
 AC QYWL0; 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Structural protein (Fragment).
 OS Hepatitis E virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;
 OC Hepatitis E-like viruses.
 OX NCBI_TaxID=12461;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=BCN.
 RX MEDLINE=99013657; PubMed=9797311;
 RA Pina S., Jofre J., Emerson S.U., Purcell R.H., Girones R.;
 RT "Characterization of a strain of infectious hepatitis E virus isolated
 RT from sewage in an area where hepatitis E is not endemic.";
 RL Appl. Environ. Microbiol. 64:4485-4488(1998).
 DR EMBL; AF058684; AAC77808.1;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR004261; SP2.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF03014; SP2; 1.

FT NON TER 605 605
SQ SEQUENCE 605 AA; 64740 MW; A088F76P874B2B72 CRC64;
Alignment Scores:
Pred. No.: 3,46e-237 Length: 605
Score: 247.00 Matches: 247
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 37.42% Indels: 0
DB: 12 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q9YWL0 (1-605)
QY 1075 AATGGTCTCGTAGATCGCGCGGATAGCCCTCACCTGTTCACCTTGCTGACACT 1134
Db 359 AsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThr 378
QY 1135 CTGCTTGGCGCCTCGCCAGCAAGATTGATTTCGTCGGCTGGTGGCCAGCTGTTCTACTCC 1194
Db 379 LeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSer 398
QY 1195 CGTCCCGTTCTCAGCAATGGCGGACGACGCTTAAGTTGTATACATCTAGAGAAT 1254
Db 399 ArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsn 418
QY 1255 GCTCAGCAGGATAAGGCTATTGCAATCCGCATGCACATTGACCTCGGAGAATCTCGGTG 1314
Db 419 AlaGlnGlnAspLysGlyIleAlaProHisAspLeuGlyGluSerArgVal 438
QY 1315 GTTATTCAGGATTATGATAACCAACATGAACAAGATCGCGCGAGCTTCTCCAGCCCA 1374
Db 439 ValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaPro 458
QY 1375 TCGCGCCTTCTCTGTCCTCGAGCTAATGATGCTGTTGGCTCTCTCCTCACCGTGCC 1434
Db 459 SerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAla 478
QY 1435 GAGTATGACCACTTATGCTTCTGACTGGCCGAGTTTATGTTTCTGACTCTGTG 1494
Db 479 GluTyrAspGlnSerThrTyrGlySerThrGlyProValTyrValSerAspSerVal 498
QY 1495 ACTTGTGTTAATGTTGACGCGCGCGAGCGCTTCCCGTCTGCTCGATTGGACCAAG 1554
Db 499 ThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLys 518
QY 1555 GTCACACTTCACGTCGCCCCCTCTCCACCATCCAGCAGTACTCGAAGACTTCTTGTTC 1614
Db 519 ValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheVal 538
QY 1615 CTGCGCGCTCCGCGTAAAGCTCTCTTCTGGGAGCGAGGCAACAATAAGCGCGGTACCCCT 1674
Db 539 LeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrPro 558
QY 1675 TATAATTATAACCACTGCTAGCAGCAACTGCTTGTGAGAAATGCGCGCGCACCGG 1734
Db 559 TyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGlyHisArg 578
QY 1735 GTCGCTATTTCCACTTACACCTAGCTGGGTGCTGGTCCGCTCTCCATTTCTGCGGTT 1794
Db 579 ValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSerAlaVal 598
QY 1795 GCGGTTTTAGCCCCCACTCT 1815
Db 599 AlaValLeuAlaProHisSer 605
RESULT 13
O91856 PRELIMINARY; PRT; 660 AA.
AC O91856;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.

OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=94-Egypt;
EX MEDLINE=99105430; PubMed=9890424;
RA Tsarev S.A., Binn L.N., Gomatos P.J., Arthur R.R., Monier M.K.,
van Cuyck-Gandre H., Longer C.F., Innis B.L.;
RT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
J. Med. Virol. 57:68-74(1999).
DR EMBL: AF0511352; AAC35764.1; -
DR GO: GO:0005398; P: structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71026 MW; 16C560FA16941F2A CRC64;
Alignment Scores:
Pred. No.: 3,23e-204 Length: 660
Score: 214.00 Matches: 214
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 32.42% Indels: 0
DB: 12 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x O91856 (1-660)
QY 1069 AGTACTAATGGTCTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTTGCT 1128
Db 357 SerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAla 376
QY 1129 GACACTCTGCTTGGCGCCTGCGCAGACAGATTGATTTCGTCGGCTGGTCCAGCTGTT 1188
Db 377 AspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPhe 396
QY 1189 TACTCCCGTCCCTGCTCTCAGCAATGGCGAGCCAGCTGTTAAGTTGTATACATCTGTA 1248
Db 397 TyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal 416
QY 1249 GAGAAATGCTCAGCAGGATAAGGCTATTGCAATCCCGCATGACATTGACCTCGGAGAACT 1308
Db 417 GluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspLeuGlyGluSer 436
QY 1309 CGTGTGTTATTTCAGGATTATGATAACCAACATGAACAAGATCGCGCAGCTTCTTCCA 1368
Db 437 ArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 456
QY 1369 GCCCCATCGCGCTTCTCTGTCCTCGAGCTAATGATGCTGTTGGCTCTCTCTCACC 1428
Db 457 AlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThr 476
QY 1429 GCTCCGAGTATGACCACTCCACTTATGGCTCTTCGACTGGCCGAGTTTATGTTTCTGAC 1488
Db 477 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 496
QY 1489 TCTGTGACCTTGTTAATGTTGCGACCGCGCGCAGCGCTTCCCGGTGCTCGATTTG 1548
Db 497 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyr 516
QY 1549 ACCAAGTCCACTTCAGCGTCCGCCCTCTCCACCATCCAGCAGTACTCGAGACCTTC 1608
Db 517 ThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhe 536
QY 1609 TTTGCTCCTCCGCTCCGCGTAAAGCTCTCTTTCTGGGAGCGAGGCAACAATAAGCCGG 1668
Db 537 PheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGly 556
QY 1669 TACCTTTAATAATTAACACCACTGCTAGCGACCAACTGCTT 1710
Db 557 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeu 570

```
RESULT 14
Q81860 PRELIMINARY; PRT; 248 AA.
ID AC Q81860;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Structural protein (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BurmaEpi-1-3;
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
ID AC Q81860;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Complete genome sequence.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Uigh179;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; D11093; BAA01867.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 248 AA; 26859 MW; 3C8139211E2D045F CRC64;

Alignment Scores:
Pred. No.: 3,75e-202 Length: 248
Score: 212.00 Matches: 212
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 32.12% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q81860 (1-248)
QY 1219 GAGCCGACTGTTAAGTTGTATACATCTGTAGAGAATCTCAGCAGGATGAAGGGTATTGCA 1278
DB 1 GluProThrValIysLeuTyrThrSerValGluAsnAlaGlnAspLysGlyIleAla 20

QY 1279 ATCCCGCATCACATTGACCTCGGAGAACTCTCGTGTGTTATTTCAGGATTATGATAACCAA 1338
DB 21 IleProHisAspLeuAspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGln 40

QY 1339 CATCAACAAGATCGCCGACGCTTCCTCCAGCCCATCGCCGCTTCTCTGCTCTTCGA 1398
DB 41 HisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeuArg 60

QY 1399 GCTTAATGATGCTTTGGCTCTCTCACCCTGCGCAGTATGACCATTTATGGC 1458
DB 61 AlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGly 80

QY 1459 TCTTCGAGTGGCCAGTTTATGTTTCTGACTCTGTGACCTTGGTAAATGTTGCGACCGGC 1518
DB 81 SerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAlaThrGly 100

QY 1519 GCGCAGGCGCTTGGCCGCTCGCTCGATTGGACCAAGGTCACTTGACGCTCGCCCTC 1578
DB 101 AlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeu 120

QY 1579 TCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTCGCGCTCCGCGGTAAAGCTCTCT 1638

RESULT 15
Q81878 PRELIMINARY; PRT; 660 AA.
ID AC Q81878;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Complete genome sequence.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Uigh179;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; D11093; BAA01867.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70937 MW; 5412313F1A2EF4A9 CRC64;

Alignment Scores:
Pred. No.: 3,14e-187 Length: 660
Score: 197.00 Matches: 197
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.85% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x Q81878 (1-660)
QY 1006 GACGGAGCTCCGAGCTCACCACGAGCTGCTACCGCTTTATGAGGACCTCTATTTT 1065
DB 336 AspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLysAspLeuTyrPhe 355

QY 1066 ACTAGTACTAATGGTGTGCTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTT 1125
DB 356 ThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeu 375

QY 1126 GCTGACACTCTGCTTGGCGGCTCGCGACAGAATTGATTTCTGCGGTGGTGGCCAGCTG 1185
DB 376 AlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeu 395

QY 1186 TTCTACTCCGCTCCGCTGCTCAGCCCAATGGCGAGCCGACTGTTAAGTTGTATACATCT 1245
DB 396 PheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSer 415

QY 1246 GTAGAGATGCTCAGCAGGATAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAA 1305
DB 416 ValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspLeuGlyGlu 435
```

```
Qy 1306 TCTCGTGTGGTTATTTCAGATTATGATAACCAACATGAACAGATCGGCCGACGCTTCT 1365
Db |||||
436 SerArgValValIleGlnAspTyrAspAsnGlnHisGlnAspArgProThrProSer 455
Qy 1366 CCAGCCCATCGCGCCCTTCTCTGTCCTTCGAGCTAATGATGTGCTTTGGCTCTCTCTC 1425
Db |||||
456 ProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeu 475
Qy 1426 ACCGCTGCCGAGTATGACCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTCT 1485
Db |||||
476 ThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSer 495
Qy 1486 GACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGCAGCGGTTGCCCGGTGCTCGAT 1545
Db |||||
496 AspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAsp 515
Qy 1546 TGGACCAAGTCCACACTTGACGGTCGCCCTCTCCACCATCCAGCAGTAC 1596
Db |||||
516 TrpThrIysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyr 532
```

Search completed: August 16, 2004, 13:55:37
Job time : 121.073 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:41:04 ; Search time 20.0151 Seconds
(without alignments)
10229.686 Million cell updates/sec

US-09-851-410a-6_COPY_5147_7129

Perfect score: 660
Sequence: 1 ATGGCCCTCGGCTATTT.....GTAAACTCGGAGTTGTAG 1983

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Word size: 1

Total number of hits satisfying chosen parameters: 663654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/uspto spool p/US09851410/runat 16082004 125718 9682/app_query.fasta_1.4238
-DB=Issued Patents AA -QMT=fastan -SUFFIX=oli.rai -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410 @CGN 1 1 45 @runat 16082004 125718 9682 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQ=0 -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEX=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	660	100.0	660	1	US-08-240-049B-15
2	660	100.0	660	1	US-08-259-148A-19
3	660	100.0	660	1	US-08-484-054-19
4	660	100.0	660	2	US-07-876-941A-19
5	660	100.0	660	3	US-08-478-507-8
6	660	100.0	660	3	US-08-542-634-13
7	660	100.0	660	3	US-09-128-275A-8
8	660	100.0	660	3	US-08-477-292-13
9	660	100.0	660	4	US-09-553-427-8
10	660	100.0	660	4	US-09-462-606-50
11	660	100.0	660	4	US-07-870-985A-19
12	660	100.0	660	5	PCT-US95-13703-13

13	549	83.2	549	3	US-08-542-634-15	Sequence 15, Appl
14	549	83.2	549	3	US-08-477-292-15	Sequence 15, Appl
15	549	83.2	549	5	PCT-US95-13703-15	Sequence 15, Appl
16	540	81.8	540	3	US-08-542-634-25	Sequence 25, Appl
17	540	81.8	540	5	PCT-US95-13703-25	Sequence 25, Appl
18	525	79.5	525	5	US-08-542-634-27	Sequence 27, Appl
19	525	79.5	525	5	PCT-US95-13703-27	Sequence 27, Appl
20	512	77.6	660	4	US-09-172-699-2	Sequence 2, Appl
21	436	66.1	436	1	US-08-259-148A-17	Sequence 17, Appl
22	436	66.1	436	1	US-08-484-054-17	Sequence 17, Appl
23	436	66.1	436	2	US-07-876-941A-17	Sequence 17, Appl
24	436	66.1	436	4	US-07-870-985A-17	Sequence 17, Appl
25	413	62.6	552	4	US-09-172-699-16	Sequence 16, Appl
26	413	62.6	561	4	US-09-172-699-20	Sequence 20, Appl
27	408	61.8	660	4	US-09-462-606-12	Sequence 12, Appl
28	396	60.0	660	4	US-09-462-606-49	Sequence 49, Appl
29	327	49.5	327	1	US-08-240-049B-13	Sequence 13, Appl
30	327	49.5	327	1	US-08-259-148A-15	Sequence 15, Appl
31	327	49.5	327	1	US-08-484-054-15	Sequence 15, Appl
32	327	49.5	327	2	US-07-876-941A-15	Sequence 15, Appl
33	327	49.5	327	3	US-08-542-634-17	Sequence 17, Appl
34	327	49.5	327	3	US-08-477-292-17	Sequence 17, Appl
35	327	49.5	327	4	US-07-870-985A-15	Sequence 15, Appl
36	327	49.5	327	5	PCT-US95-13703-17	Sequence 17, Appl
37	321	48.6	660	3	US-08-840-316-2	Sequence 2, Appl
38	321	48.6	660	3	US-08-809-523-2	Sequence 2, Appl
39	321	48.6	660	3	US-08-471-971-2	Sequence 2, Appl
40	321	48.6	660	4	US-09-402-776-2	Sequence 2, Appl
41	321	48.6	660	5	PCT-US93-08849A-2	Sequence 2, Appl
42	321	48.6	660	5	PCT-US93-08849-2	Sequence 2, Appl
43	284	43.0	660	4	US-09-462-606-56	Sequence 56, Appl
44	264	40.0	660	4	US-09-462-606-53	Sequence 53, Appl
45	259	39.2	660	4	US-09-462-606-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1

US-08-240-049B-15
; Sequence 15, Application US/08240049B
; Patent No. 5686239
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Tam, Albert W.
; APPLICANT: Varbough, Patrice O.
; TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,049B
; FILING DATE: 09-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles K. Sholtz
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
; INDIVIDUAL ISOLATE: ORF-2
; US-08-240-049B-15

Alignment Scores:

Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-08-240-049B-15 (1-660)

```
QY 1 ATCGCGCCCTCGGCTATTGTTGCTGCTCTCATGTTTGTGCTATGCTGCGCGCCCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCGGTCAAGCTGCGCGCGTCTGCGCGTCTGCGCGCGCGCGCGCGTTCGCGGT 120
Db 21 ProProGlnProSerGlyArgArgGlyArgGlySerGlySerGlyGlyGly 40
QY 121 TCTCGGGTCACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTrpIleHisProThrAsn 60
QY 181 CCTTCGCCCCCGATGTCACCGTGGCGCGCGCTGGAGCTGCTGTTCCCAACCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCCGCTTGGCTGACCGGCGCGCGCGCGCGCGCGCTCACGTGCT 300
Db 81 ArgProLeuGlySerAlaTrpArgGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACCTACCAAGCTGGGCGCGCGCTAACCGCGTCTCGCGCGCGCGCGCGCGCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CGAGTCCGCTGATCGACTCCCGCGCGCATCTTCGCGCGCGAGTATACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTrpAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTrpAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCCCTTCAGGACGGCACCAATACCATATATATGCGCGCGAGCTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATATGCCAGTACCGGTTGCCGTGCCAATCCGTTACCGCGCGTGTGCCCAAT 600
Db 181 AsnTyrAlaGlnTrpArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGGTTTACGCCATCTCCATCTCATCTGCGCGCACAGCACACCGCGGAG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATGATCAATTAACCTCGAGGATGTTTCGTATTATTAGTCCAGCGCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTGAGCTGTGATCCCAAGTAGCGCCTACACTATCGTAACCAAGGCTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGTTGCTGAGGAGGAGGCTACTCTGCTGTTGTTATGTTGCATA 840
Db
```

```
Db 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTGTAATTCCTATACACCTATACCGGTGCGCTCGGGCTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTCAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTGCGGTGCGGACGAGACTGCCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACGCGCTGCTACCCGCTTATGAAGACCTCTATTTTACTAGTACTAATGGT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTTGCTGACACTCTGCTT 1140
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCCAGACAATTGATTTCGTGCGGTGGTGGCCAGCTGTTCCTACTCCCGTCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATCGCGAGCGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGAGAATCTCGTGTGTTATT 1320
Db 421 GlnAspLysGlyIleAlaIleProHisAspLleAspLeuGlyGluSerArgValValIle 440
QY 1321 CAGGATTATGATAACCAACATGAACAGATCGCGCGCGCTTCTCAGCCCATCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTCTCGAGCTAATGATGTCTGCTGCTCTCTCAGCGGTGCGGATAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCCACTTATGGCTCTTCGACTGGCGCCAGTTATGTTTCTGACTCTGTGACTTG 1500
Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAAATGTTGCGACCGCGCGCGCGCGCTTGGCCGCTCGATTCGATTCGACAGGTCA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTGACGTCGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCGCG 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPheValLeuPro 540
QY 1621 CTCGCGGTAAAGCTCTCTTTCGGGAGCGGACCAACATAAAGCCGGGTACCTTATAAT 1680
Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGTAGCGACCAACTGCTGTGCGAAGATGCCCGCGGACCGCGGTGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTCGCTAGCTATGCTTGAGGATACCTTGAGACTTACCTCCCGCGCGC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGATGATTCTGCGCACAGTGCAGCGCGCTTGGCTTTCAGGGCTGCGCTTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
```

QY 1921 CAGTCTACTGCTGAGCTTCACGGCTTAAGCTGAAGCTGGCTAAACTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 2

US-08-259-148A-19
Sequence 19, Application US/08259148A
Patent No. 5741490
GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
APPLICANT: Bradley, Daniel W.
APPLICANT: Twu, Jr-Shin
APPLICANT: Purdy, Michael A.
APPLICANT: Tam, Albert W.
APPLICANT: Krawczynski, Krzysztof Z.
APPLICANT: Yarbough, Patrice D.
TITLE OF INVENTION: Hepatitis B Virus Vaccine and Method
NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/259,148A
FILING DATE: 13-JUN-1994
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APR-1989

REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.20
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO

ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
US-08-259-148A-19

Alignment Scores: 0 Length: 660
Pred. No.:

Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-08-259-148A-19 (1-660)
QY 1 ATGGCGCCCTGGCGCTATTTTGTCTCTCTCATGTGTTTTTGTCTATGTCTCCCGCGCCA 60
Db 1 MetArgProArgProLleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGGTTCAGCGCTCTGGCGCGCTGTGGGGGGCGACGCGGGTTCGGGGGTGGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40
QY 121 TTTCTGGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCGATGTCCACCGCTGGCGCGGGGTGGACCTCGTGTTCGGCAACCCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTTGGCGTGACCGACCGCCCGCGCGCTGGTTCCTCAGTCTGT 300
Db 81 ArgProLeuGlySerAlaTipArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACTTACCACAGCTGGGGCGCGCTTAACCGGGTTCGGCGCGCCCATGACACCCCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTGCGACTCCCGCGCGCCCATCTTGGCGCGGCAGTATAACCTATCAAC 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTTACCTCTTCCTGGCGCACCGGCACTAACCTGGTGTCTTTATGCCCGCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGGCTTTTACCTCTCAGGACCGCACCATACCATATATATATGCGCACCGAAGTCTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCCGCTACCGGTTGCCCGTGCACAAATCCGTTTACCGCGCGCTGGTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTGGCGGTTACGCCATCTCCATCTCATCTCTGGCCACAGACACACACCGCGAGC 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCAATACCTCGAGGATGTTCGTATTATTAGTCAGCGCGGCATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCAAGTAGCGCTACACTATCTGTAACCAAGCTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGTACTCTGGTCTTGTATTATGCTTTGCAATA 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTACTCGTAAATTCCTATATAACCTATACACCTATACCGGTGCGCTCGGGCTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAACCAATACCGGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCCAGCACTGCTCGCCACCGCTTCTGTGGGGTGGCGAGCGGACTGCCGAG 1020

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/876,941A
FILING DATE: 01-MAY-1992
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 505,889
FILING DATE: 05-APRIL-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.33
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
US-07-876-941A-19

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-07-876-941A-19 (1-660)

QY 1 ATGCGCCCTCGGCTATTTGTTGCTCTCCTCATGTTTTCCTATGCTGCGCGGCA 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 CGCGCCGCTCAGCCGCTCTGCGCGCGCGCTGCTGCGCGCGCGAGCGGCTTCGCGGTGT 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 21 ProProGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlySerGlyGly 40
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 TTCTGGGGTGACGGGTGATTTCTCAGCCCTTCGCAATCCCTATATTATCCAAACCAAC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrlleHisProThrAsn 60
QY 181 CCCTTCGCCCGCGATGTACCGCTGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTGCGGTGACAGGCCCGCGCGCTGCTCAGTGT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGCACTACCAACAGCTGGCGCGCGCTAACCGCGGTGCTCGCGGCCATGACACCCCG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCTGATGTGCACTCCCGCGCGCCATCTTGGCGCGGCACTATAACTATCAACA 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 ProValProAspValAspSerArgGlyAlaIleleuargGlnTyraAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCCTGCGGCGCGCACTAACCTGTCTTTTATGCGGCCCTCTT 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyraAlaProLeu 160
QY 481 AGTCCGCTTTTACCCCTTCAGGACGCGCAACATACCATATATATGCGCGGAGCTTCT 540
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGCTTGGCGGTGCGCACCAATCCGTTACCGCGCGCTGGTCCCAAT 600
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 AsnTyraAlaGlnTyraArgValAlaAlaArgAlaThrIleargTyraArgProLeuValProAsn 200
QY 601 GCTGTGGCGGTACGCCATCTCCATCTANTCTGGCCACAGACACCAACCCCGGACG 660
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 201 AlaValGlyGlyTyraAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCMAATACTCGACGATGTTGCGTATTTAGTCCAGCGCGCAT 720
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGCTCGCGCTTC 780
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyraArgAsnGlnGlyTrpArgSer 260
QY 781 GTGAGACCTCTGGGTGCTGAGGAGGAGGCTACCTCTGGTCTTGTATTGCTTTGCATA 840
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCACTCGTAAATTCCTATATAACACCTATACCGCTCGGCTCGGCTGTTG 900
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 281 HisGlySerLeuValAsnSerTyraAsnThrProTyraThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGAAACCTTACCCCGGTAACCAATAACGCGGTC 960
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCGTCGGGTGCGGACGGGACTGCGGAG 1020
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 321 SerArgTyraSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCAACGCTGCTACCCGCTTATGAAGGACCTCTATTTTACTAGTACTAATGTT 1080
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 341 LeuThrThrThrAlaAlaThrArgPheMetIysAspLeuTyraPheThrSerThrAsnGly 360
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTGTGTTCAACCTTGTCTGACACTCTGCTT 1140
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GSCGCGCTCGCGACAGAATTGATTTGCGGTGCTGCGGTGCGGCTGCTTCTATCCGCTGCC 1200
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyraSerArgPro 400
QY 1201 GTTGTCTCAGCAATGCGGAGCCGCTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 401 ValValSerAlaAsnGlyGluProThrValIysLeuTyraThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAGGGTATTGCAATCCCGCATGACATGACATGACGAGAAATCTCGTGTGTTATT 1320
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440

```
QY 1321 CAGGATTATGATAACCAACATGAAGATCGCGCAGCGCTTCTCCAGCCCATCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGTCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCCACTATGCTCTTCTGACTGCGCCAGTTTATGTTTCTGACTCTGTGACTTGG 1500
Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGCGACCGCGCGAGCGGCTTGGCCCGTTCGCTGATTGGACCAAGGTCACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspThrThrValThr 520
QY 1561 CTTGACGGTGGCCCGCTCTCCACATCCAGCAGTACTCGAAGACCTTTCTTGTCTGCGCG 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAAGCTCTCTTCTGGGAGGCGAGCAGCAGCACTAAAGCCGGGTACCTTATAAT 1680
Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAATGCGCGCGCAGCGGTCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTCCACTTACACCACTAGCTGGTGTGCTGCTGCTGCCATTTCTCCATTTCTCGGTTGCGGT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCATCTGCGGTAGCTTGTGAGGATACCTTGACTACCTGCGCGCGCGC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCTTCAGGCTGCGGCTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTGCTGAGCTTCCAGCGCTTAAAGATGAAGGTGGTAAACTCGGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660
```

RESULT 5

```
US-08-478-507-8
; Sequence 8, Application US/08478507
; Patent No. 6120988
; GENERAL INFORMATION:
; APPLICANT: Reves, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fy, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6120988-A/No. 6120988-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507
```

FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/279,823
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,078
FILING DATE: 05-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0183.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-507-8

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-08-478-507-8 (1-660)

QY 1 ATGCGCCCTCGGCCTATTTTGTGTGCTCTCATGTTTTCCTATGCTGCCGCGCCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGGTACGCTGTGGCGCGCTGTGGCGCGCGCGCGCGGTTCGCGCGTGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40
QY 121 TTCTGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTATCATCAACAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCCGATGTCCACGCTCGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTTGGCGTGACGAGGCCCGCGCGCTTCCTTCAGTGTGT 300
Db 81 ArgProLeuGlySerAlaTyrArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTACACAGCTGGGGCGCGCGCTTAACCGCGGTGCTCCGCGCCATGACACCCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTGCTCCCGCGCGCCATCTTCCCGCGCGAGTATTAACCTATCA 420
Db 420

Db 121 ProValProAspValAlaAspSerArgGlyAlaAlaLeuArgArgGlnTyrAsnLeuSerThr 140
Qy 421 TCTCCCTTACCTCTTCGCGGCGCCACCGGCACTAACTGGTCTCTTATGCGCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
Qy 481 AGTCCCGCTTTACCTCTCAGGACGGCACCAATACCCATATATGGCCACGGAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisLeuMetAlaThrGluAlaSer 180
Qy 541 AATTATGCCAGTACCGGGTTGCCGTCGACCAATCCGCTTACCGCGCTGGTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
Qy 601 GCTGTCCGGCGGTACGCATCTCCATCTCTTGGCCACAGACCAACCCCGGACG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr 220
Qy 661 TCCGTTGATATGAATTCATAAACCCTCAGCGATGTTCGTATTTTAGTCCAGCCGGCATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
Qy 721 GCTCTGAGCTGTGATCCCAAGAGGCGCTACACTATCGTAAACCAAGCGTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
Qy 781 GTCGAGACCTCGGGGCTGAGGAGGAGCTACCTCTGCTCTGTATGCTTTGCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
Qy 841 CATGGCTCACTCGTAAATTCCTATPACTAATACACCTTATACCGGTGCCCTCGGCTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
Qy 901 GACTTTCCTTACCTTGGTGTAGTTTCGCAACCTTACCCCGGTAAACCAATACCGCGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
Qy 961 TCCGTTATTCACGACTGCTCGCACCGCTCGTGGTGGGACGGGACTCGCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
Qy 1021 CTCACCAACCGGCTGCTACCGCTTTATGAAGGACCTCTATTTACTAGTACTAATGGT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
Qy 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTGTGTGACACTCTGCTT 1140
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
Qy 1141 GCGGCGCTGCGACAGAAATTCATTTGCTGGCTGGTGGCCAGCTGTCTACTCCCGTCCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
Qy 1201 GTTCTCTCAGCAATGCGGACCGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
Qy 1261 CAGATAAGGTATTCATCCCGCATGCCATGACATTCACCTCGGAGATCTCGTGTGTTATT 1320
Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
Qy 1321 CAGATTATGATAACCAATGAACAGATCGCGGACGCTTCTCCAGCCCGCCATCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
Qy 1381 CTTTCTCTGTCTCGAGCTTAATGATGTCTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuThrPheLeuSerLeuThrAlaAlaGluTyr 480
Qy 1441 GACAGTCCACTTATGGCTCTCGACTGGCCAGTTATGTTCTGACTCTGTGACCTTG 1500
Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500

Qy 1501 GTTAATGTTTCGACCGCGCGCGCGGCGTTGCCGCTCGCTCGATTGGACCAAGTCCACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLysValThr 520
Qy 1561 CTTGAGCGTTCGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTGTCTGCGCG 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
Qy 1621 CTCGGGGTAAGCTCTCTTTCGGAGGCGAGGACAACTAAAGCCGGGTACCTTATAAT 1680
Db 541 LeuArgGlyLysLeuSerPheTyrPheGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
Qy 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTTCGAGAATCCCGCGGCGACCGGTCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
Qy 1741 ATTTCCACTTACACCACTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
Qy 1801 TTAGCCCCCAGCTCTGCGTAGCATTTGCTTGGAGTACCTTGGACTACCTGCGCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
Qy 1861 CATACTTTTGATGATTTCTCCCGAGAGTCCCGCCCGCTTGGCTTCAGGCGTGGCTTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
Qy 1921 CAGTCTACTGCTGAGCTTACCGGCTTAAAGTGAAGTGGGTAAACTCGGAGCTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 6

US-08-542-634-13
; Sequence 13, Application US/08542634
; Patent No. 6214970
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; APPLICANT: McAttee, C. Patrick
; APPLICANT: Yarbough, Patrice O.
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542,634
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein


```
; HYPOTHETICAL: NO
; ORIGINAL SOURCE: Hepatitis E Virus (Burma strain)
; INDIVIDUAL ISOLATE: ORF-2
; INDIVIDUAL ISOLATE: US-08-542-634-13
US-09-851-410a-6_COPY_5147_7129 (1-1983) x US-08-542-634-13 (1-660)

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-851-410a-6_COPY_5147_7129 (1-1983) x US-08-542-634-13 (1-660)
QY 1 ATGCGCCCTCGCCCTATTTGTTGCTCTCTCATGTTTTTGCCTATGCTGCGCCGCCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCCCGGTACCGCTCTGCGCGCGCTGCTGGCGCGCGAGCGGCTTCCGGCGGTGT 120
DB 21 ProProGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlySerGlyGly 40
QY 121 TTCTGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCGATGTCAACCGCTCGCGCGGTGAGCCTCTGTTGCGCAACCCGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTGCGGTGCGTACCAGCGCCAGCGCGCGCTGCGCTCAGCTGT 300
DB 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACCTACACAGCTGGGGCGCGCGCTAAACCGCGGTGCTCGCGCCCATACACACCGG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTCGACTCCCGCGCGCCATCTTTGCGCGCGCAGTATCAACCTATCA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnFyrAsnLeuSerThr 140
QY 421 TCTCCCTTACTCTTCGCTGCGCCACCGCCTAACCTGCTCTTATGCGCGCCCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCCCTTCAGACCGCACCAATACCATATATATGCGCACGGAAGCTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGTTCGCCGTGCCAATCCGTTACCGCGCGTGTGCCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTGCGCGGTAGCCATCTCCATCTCATTTCTGCGCCACAGCACACCGCCGCGAG 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
QY 661 TCCGTTGATATCAATTCAATACCTCGACGGATGTTGCTATTTAGTCCAGCGCCGCAT 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCAAGTAGCGCCTACACTATCGTAACCAAGCTGGCGCTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGCTACCTCTGGTCTGTTATGCTTTCGAT 840
DB 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAAATTCCTATATACTAATACACCCCTATACCGGTGCTCGGGCTGTG 900
```

```
DB 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTTAACACCAATACCGGGTTC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGTTATTTCAGCACCTGCTCGCCACCGCTTCGTCGCGGTGCGAGCGGACTGCCGAG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGCTGCTACCGCTTTTATGAAGACCTCTATTATTTACTAGTACTAATGT 1080
DB 341 LeuThrThrThrAlaAlaThrArgPheMetIysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCCCTGTTCAACCTTGCTGACACTCTGCTT 1140
DB 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCGCACAGAATTGATTTTCGTCGCGTGGTGGCCAGCTGTTCTACTCCCGTCCC 1200
DB 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
DB 401 ValValSerAlaAsnGlyGluProThrValIysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGTAAGGTATGCAATCCCGCATGACATGACCTCGGAGAAATCTCGTGTGTTATT 1320
DB 421 GlnAspIysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
QY 1321 CAGGATATATGATAACCAACATGAACATCGCGCGAGCCCTTCTCCAGCCCCATCGCGC 1380
DB 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTTCTCTGCTCTCGAGCTAATGATGTGCTTGGCTCTCTCTCACCGCTGCCAGTAT 1440
DB 461 ProPheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACAGTCACTATGCTTCTCGACTGCGCCAGTTATGTTTCTGACTCTGTGACCTTG 1500
DB 481 AspGlnSerThrTyrGlySerSerGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGGACCGCGCGCAGCGCTGTCGCGTCTCGATTTGACCAAGGTCA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrIysValThr 520
QY 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTCTGCTGCG 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAGCTCTCTTCTGGGAGGAGGAGCACAACCTAAAGCCGGTACCTTATAT 1680
DB 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGCGAGAATGCGCGCGGACCGGTCGCT 1740
DB 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACCACTAGCCTGGGTGCTGCTGCCCTTCCATTTCTCGGTTGCCGTT 1800
DB 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGTTGAGGATACCTTGGACTACCTGCCCGCGCC 1860
DB 601 LeuAlaProHisSerAlaLeuLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGTATTTCTGCCACAGTGCGCCCTTGGCTTTCAGGGCTGCGCTTC 1920
DB 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCTGCTGAGCTTCAGCGCTTAAAGATGAAGGTGGTAAAGCTCGGAGTTG 1980
```

Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

```
RESULT 7
US-09-128-275A-8
; Sequence 8, Application US/09128275A
; Patent No. 6229005
; GENERAL INFORMATION:
; APPLICANT: Reves, Gregory R
; APPLICANT: Yarborough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fiv, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6229005-A/No. 6229005-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,275A
; FILING DATE: 03-AUG-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-128-275A-8
```

Alignment Scores:
Pred. No.: 0
Score: 660.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Length: 660
Matches: 660
Conservative: 0
Mismatch: 0
Indels: 0

```
DB: 3 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-128-275A-8 (1-660)
QY 1 ATCGCGCCTCGGCTATTTTGTCTCTCTCTCATGTTTGTCTATGTCGCGCGCCCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCGCGGTCAGCCGCTCTGGCCCGCTGCGGCGCGCGAGCGCGGTTCCGCGGCGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgGlySerGlyGlyGlyGly 40
QY 121 TTCTGGGGTGACCGGGTTCGATCTCTACCGCTTCGCAATCCCTATATTCATCAACA 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCCCGATGTCACCGCTGCGGCGCGGCTGGACCTCGTGTTCGCCAACCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCCCGCTTGGCGTGACCGCGCGCGCGCGCGCGCGCTGCTGCTCAGTC 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTACCACTCGGCGCGCGCTAACCGCGCTGCTCGCGCGCGCGCGCGCGCGCGCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTCCTGATGTCGACTCCCGCGCGCGCATCTTCGCGCGCGAGTATACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCTTTTACCCCTTCAGGACGCGCACCAATACCATATATATGCGCGCGAGGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGCGGGTACGCCATCTCCATCTCATCTGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATTAACCTCGAGCGGATGTTGCTATTTTGTCTAGTCCAGCG 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCAAGTACCGCGCTTACACTATCGTAACCAAGCTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCATCTGTAATTCCTATATAACCTATACACCTATACCGGCTGCGCGCGCTGGT 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGCGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCAGGACTGCTCGCGCGCGCTTCTGCGCGGTCGCGAGCGCGGTCGCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCACCGCGCTGCTACCGCGCTTTTATGAAGGACCTCTATTTTACTAGTACTATGGT 1080
```


361 CCAGTGCCTGATGTGACATCCCGCGCGCACTCTTGGCGCGGAGTATACCTATCAACA 420
Db
121 ProValProAspValAspSerArgGlyAlaLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCGCTTACCTCTTCCGTGGCCACCGGACCTAACCTGGTCTCTTTATGCGCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTTCCCTCCAGACGCGCACCAATACCATATATATGCGCGGAGGAGTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGCTTCCCGTGGCCACCAATCCGTTACCGCCCGCTGGTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCCGGGTTAGCCATCTCCATCTCATCTGCTGGCCACAGACACACACCGCGAGC 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATAACCTCGACGGATGTTCTGTATTTAGTCCAGCCCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGGAGCGCTACACATATCGTAACCAAGGCTGGCGCTC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
QY 781 GTCGAGACCTCTGGGTGGCTGAGGAGGAGTACCTCTGGTCTGTGTATGCTTTGCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAAATTCCTATATACTATACACCTATACCGGTGCGCTCGGGTGTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTGGCTTGGAGCTTGAGTTTCGAACCTTACCCCGGTAAACCAATACCGGGTTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTATTCCAGACACTGCTCGCCACCGCTTGGTGGGTGGCGGAGGAGTCCGAG 1020
Db 321 SerArgTyrSerSerThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrAsnGly 340
QY 1021 CTCACACACCGCTGCTACCGCTTTATGAAGGACCTCTATTCTAGTACTAATGTT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrAsnGly 360
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTT 1140
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGGACAGAAATGATTTCGTCGGCTGGCGAGCTGTTCTACTCCGCTCCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAGGTTATGCAATCCCGCATGACATGACCTCGGAGATCTGCTGTGTTAT 1320
Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 440
QY 1321 CAGGATATGATAACCAATGAAACAGATCGGCGAGCGCTTCTCCAGCCCATCGCG 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTCCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTyrPheSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCAGTCCACTATGGCTCTTCGACTGCGCCAGTTATGTTCTGACTCTGTGACCTTG 1500

481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTATGTTGACACCGCGCGAGCGGCTTGGCCGCTGCTCGATTGGACCAAGTCA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTyrThrLysValThr 520
QY 1561 CTTGACGGTCCGCTCTCCACCATCCAGCAGTACTCTCAAGACCTTCTTTGTCCTGCG 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAACTCTCTTTTGGAGGCGAGCACAACTAAAGCCGGGTACCTTATAAT 1680
Db 541 LeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGCTAGCGACCACTGCTTGTGAGAAATGCCCGCGGACCGGTCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCACCTTACACCACTAGCCTGGTGTGCTGCCGTCTCCATTTCTCGGTTGCCGTT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATGCTTGGAGTACCTTGGACTACCTTCCCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTGATGATTTCTGCCAGAGTCCCGCCCTTGGCTTGGCTTCCAGGCTCGCTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTGAGCTTCCAGCGCTTAAAGATGAAGGTGGTAAAACTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 9
US-09-553-427-8
; Sequence 8, Application US/09553427
; Patent No. 6379891
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; APPLICANT: Yarbough, Patrice O
; APPLICANT: Bradley, Daniel W
; APPLICANT: Krawczynski, Krzysztof Z
; APPLICANT: Tam, Albert
; APPLICANT: Fy, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; TITLE OF INVENTION: No. 6379891-A/No. 6379891-B Hepatitis Viral Agent
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/553,427
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,507
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/279,823
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/505,888
FILING DATE: 05-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/420,921
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/367,486
FILING DATE: 16-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336,672
FILING DATE: 11-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/208,997
FILING DATE: 17-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0183.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-553-427-8

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-851-410a-6_COPY_5147_7129 (1-1983) x US-09-553-427-8 (1-660)

QY 1 ATGCGCCCTCGCCGCTATTTGTGTCTCTCATGTTTTCCTATGTCGCCGCGGCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCCCGGTGACGCTCGCCGCGTGTGGCGCGGCGAGCGGGTTCGCGCGGTGT 120
DB 21 ProProGlyGlnProSerGlyArgArgGlyArgGlySerGlyGlyGly 40
QY 121 TTCTGGGTGACCGGGTGTGATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
DB 41 PheTTPGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTGGCCCGATGTACCGCTGGCGGGGCTGGACCTCGTGTTCGCAACCGGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGGCTGGGTGACAGGCCCGCGCGCGCTTCCTCAGTCGT 300
DB 81 ArgProLeuGlySerAlaTTPargAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTTACCACAGCTGGGCGCGCGCTTAACCGCGGTCTCCGCGCCCATGACACCCCG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisPaspThrPro 120
QY 361 CCAGTGCCTGATGTGACCTCGCGCGCGCATCTTGGCGCGCGCAGTATTAACCTATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTTTCGTCGCGCACCGGCACTTAACCTGTTCTTTATGCGCCCTCTT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCCGCTTTTACCCCTTCAGGACGCGCACCAATACCATATAATGCGCACGGAAGCTTCT 540

DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTGCCGTCGCCCAATCCGTTACCGCCGCTGTCGCCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCCGCGGTAGCCCATCTCCATCTCATCTCTGGCCACAGACCACCAACCCGCGAG 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTTPProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATTAACCTCGACGAGTGTCTGATTTAGTCCAGCCCGCAT 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTGTGAGCTTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGTGGCGCTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrPargSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGGCTACTCTGGTCTTGTATGCTTTGCATA 840
DB 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAAATTCCTATATACTAATACCTATACCGGTGCGGCTCGGCTGTG 900
DB 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGTC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCTTCCTCGCGGTGCGGAGCGGACTGCCGAG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACAGCGCTGCTACCGGCTTTATGAGAGGACCTCTATTTACTAGTACTAATGT 1080
DB 341 LeuThrThrThrAlaAlaThrArgPheMetIysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 CTCGCTGAGATCGCGCGGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
DB 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCCACAGAATTGATTTCTGCTGGGTGGCGAGCTGTTCTACTCCCGTCCC 1200
DB 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGCGGAGCGGACTGTAAAGTGTATACATCTGTAGAGAACTCTCAG 1260
DB 401 ValValSerAlaAsnGlyGluProThrValIysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGTATTGCAATCCCGCATGACATTCGAGAGATCTCGTGGTATT 1320
DB 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
QY 1321 CAGGATTATGATAACAACATGAACAAGATCGCGCGAGCTTCTCCAGCCCATCGCGC 1380
DB 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGTCCTTCGAGCTAATGATGTCTTTGGCTCTCTCTCACCGCTCGCGAGTAT 1440
DB 461 ProPheSerValLeuArgAlaAsnAspValLeuTTPLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACGAGTCCACTTATGGCTCTTCGACTGCGCGAGTTTATGTTCTGACTCTGACCTG 1500
DB 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGACCGCGCGGCGAGCGGCTGGCCGCTCGCTCGATTGGACCAAGGTCA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTTPThrLysValThr 520
QY 1561 CTTGACGCTCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTCTCGCGG 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrThrPhePheValLeuPro 540

QY 1321 CAGGATTATGATAACCAACATCAAGATCGCGGCGCTTCTCCAGCCCATCGCG 1380
DB 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProAlaProSerArg 460
QY 1381 CCTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTCTCTCTCACCGCTGCGAGTAT 1440
DB 461 ProPheSerValLeuArgAlaAsnAspValLeuTriLeuSerLeuThrAlaGluTyr 480
QY 1441 GACCACTCCACTATGCTCTTCGACTCGCGGCGGCTTATGTTTCTGACTCTGACCTTG 1500
DB 481 AspGlnSerThrTyrglySerThrGlyProValTyValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGCGACCGCGCGAGGCGCTTGGCGGCTCGCTGATTGGACCAAGGTCACA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTGACGGTCGCGCCCTCTCCACCTCCAGCAGTACTCGAAGACCTTTCTGCTGCGCG 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAAAGCTCTCTTCTGGGAGGCGAGCAGCACTAAAGCGGGTACCTTATAT 1680
DB 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGAATGCGCGGCGACCGGGTCGCT 1740
DB 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaGlyHisArgValAla 580
QY 1741 ATTCCACTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB 581 IleSerThrTyThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACCTTTGATGATTCTGCCAGAGTCCCGCGGCTTGGCTTCCAGGGCTGCGCTTTC 1920
DB 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
DB 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 11

US-07-870-985A-19
; Sequence 19, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07870,985A
; FILING DATE: 20-APRIL-1992

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0093.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
US-07-870-985A-19

Alignment Scores:

Pred. No.:	0	Length:	660
Score:	660.00	Matches:	660
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-07-870-985A-19 (1-660)

QY	1	ATGCGCCCTCGGCCTATTTTGTGCTCTCTCATGTTTTCCTATGCTGCCGCGCA	60
DB	1	MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20
QY	61	CGCGCGGTCACCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGTGT	120
DB	21	ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlyGlyGlyGly	40
QY	121	TTCTCGGGTGACCGGGTGGATTCTCAGCCCTTCGCAATCCCTATATTCATCAAC	180
DB	41	PheTrpGlyAspValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn	60
QY	181	CCCTTCGCGCCCGATGTTCACCGCTCGCGCGCGGCTGGACCTCTGTTGCGCAAC	240
DB	61	ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla	80
QY	241	CGACCACTCGGCTCGCTTGGCGTGACGAGGCCCGCGCGCGCGCGCGCTTCACGTGT	300
DB	81	ArgProLeuGlySerAlaTrpArgGlnAlaGlnArgProAlaValAlaSerArg	100
QY	301	AGACCTACACAGCTGGCGCGCGCGCTTAACCGGGTCTCCCGCGCGCGCGCGCG	360
DB	101	ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro	120
QY	361	CCAGTGCCTGATGTCGACTCCCGCGCGCGCATCTTTCGCGCGCGCGCTATTAAC	420

Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGTCGCGCCACCGCCTACCTGCTTCTTATGCGCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCTTTTACCCCTTCAGACGGCCACCAATACCATATATATGCGCCACGGAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGTTGCCGTCGCAATCCGTTACCGCCCTGCTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGCGGGTTAGCCATCTCCATCTCATCTCTGCGCCACACACACCGCCGAGG 660
Db 201 AlaValGlyTyrAlaIleSerIleSerPheThrProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATAACCTCGACGGATGTTGCTATTTAGTCCAGCCCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGATCCCAAGTACGACCTTACCTATCGTAACCAAGCGTGGCGTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 CTCGAGACCTCTGGGTCGCTCAGGAGGAGCTACCTCTGCTGTTGTTATGCTTGCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCCTCCTGTAATTCCTATATAATACACCTTATACCGTGCCTCGCGGTGTTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTGGCTTGGCTGAGCTTGCAGCTTCCGACCTTACCCCGGTAAACCAATACCGGGTFC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCCGCTTATCCAGACCTCTCGCCACCGCTTCGTCGCGGTCGCGGAGGAGTCCCGAG 1020
Db 321 SerArgTyrSerSerThrAlaAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGCTGCTACCCGCTTTATGAGGACCTCTATTTACTAGTACTAATGCT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTTCCACCTGTTCAACCTGCTGACACTGCTT 1140
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCCACAGAATTGATTCGTCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCCAATGGCGAGCGACTGTAAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGATTTGCAATCCGATGACATGACCTCGGAGATCTCGTGTGTTAT 1320
Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
QY 1321 CAGGATTTATGAACCAATCAACATCGGCGAGCGCTTCTCCAGCCCGCTCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTTCGAGCTAATGATGCTTGGCTCTCTCTCACCGCTCCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCCACTATGCTCTCGACTGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Db 481 AspGlnSerThrTyrGlySerThrGlyProValTyrValSerAspSerValThrLeu 500

QY 1501 GTTAATGTTGCGACCGCGCGCAGGCGCTTGCCTCGCTCGCTGATTGACCAAGTTCACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspThrThrLysValThr 520
QY 1561 CTTGACGCTGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCCTGCGG 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAGCTCTTCTTGGAGCAGGACCACTAAAGCGGCTACCTTATAT 1680
Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCATCTAGCAGCACTGTTGTCGAGAATGCCCGCGGACCGGGTCCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCTGCTTGGAGTACTTGGACTACCTGCGCTGCCCGGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTGTGATGTTCTGCGCAGAGTCCCGCCCTTGGCTTCCAGGCTGCGCTTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTGAGCTTACGCGCTTAAAGATGAAGGTGGGTAAACCTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660

RESULT 12

PCT-US95-13703-13
; Sequence 13, Application PC/TUS9513703
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND
; TITLE OF INVENTION: USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13703
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)

INDIVIDUAL ISOLATE: ORF-2
PCT-US95-13703-13

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x PCT-US95-13703-13 (1-660)

QY	1	ATGCGCCCTCGGCTATTTTGTGCTGCTCTCATGTTTTCCTATGCTGCGCGGCCA	60
DB	1	MetArgProArgProLeuLeuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro	20
QY	61	CGCCCGGTCAGCGCTGCTGCGCCCGCTGCGGCGCGCAGCGCGGTTCGGCGGTGGT	120
DB	21	ProProGlyGlnProSerGlyArgArgArgGlyArgArgSerGlyGlyGlyGly	40
QY	121	TTCTGGGTGACGGGTGATTCTCAGCCCTTCGCAATCCCTATTATTCATCCAAACAC	180
DB	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTrpIleHisProThrAsn	60
QY	181	CCCTTCGCCCCGATGTCACCGCTGCGCGCGGCGGCTGAGCCCTCGTTCGCCAACCCGCG	240
DB	61	ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla	80
QY	241	CGACCACTCGGCTCCGCTTGGCGTGACCGCCAGCGCCGCGCGCTTGCCTCACGTCGT	300
DB	81	ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg	100
QY	301	AGACCTACACAGCTGGGCGCGCGCTTAACCGCGGTGCTCCGCGCCATGACACCCCG	360
DB	101	ArgProThrThralaGlyAlaAlaProLeuThrAlaValaProAlaHisAspThrPro	120
QY	361	CCAGTGCCTTGATCGACTCCGCGCGCGCTTCGCGCGGCGAGTAACTATCAACA	420
DB	121	ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTrpAsnLeuSerThr	140
QY	421	TCTCCGCTTACCTCTTCGCTGGCCACCGCCTAACCTGCTTCTTATGCGCGCCCTCTT	480
DB	141	SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTrpAlaAlaProLeu	160
QY	481	AGTCCGCTTTACCTTCAGAGCGGACCAATACCCATATAATGCGCACGGAAGTTCT	540
DB	161	SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180
QY	541	AATTATGCCAGTACCGGCTGCGCGCCACCAATCCGTTACCGCCGCTGTCGCCCAAT	600
DB	181	AsnTrpAlaGlnTrpArgValAlaAlaAlaThrIleArgTrpArgProLeuValProAsn	200
QY	601	GCTGTCGGCGGTTACCCATCTCCATCTCATCTTGGCCACAGACCACCCCGGACG	660
DB	201	AlaValGlyGlyTrpAlaIleSerIleSerPheTrpProGlnThrThrThrProThr	220
QY	661	TCCGTTGATGATTAATCAACTCGACGGATGTTGCTATTATTAGTCAGCCCGCGCATA	720
DB	221	SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle	240
QY	721	GCCTCTGAGCTTGATCCCAAGTAGCGCTACACTATCTGTAACCAAGGCTGCGGCTCC	780
DB	241	AlaSerGluLeuValIleProSerGluArgLeuHisTrpArgAsnGlnGlyTrpArgSer	260
QY	781	GTGAGACCTCTGGGTGGCTGAGGAGGCTACCTCTGGTCTTGTTATGCTTTTCATA	840
DB	261	ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280
QY	841	CATGGCTCACTCGTAATTCCTATACCTATACACCTTATACCGGTGCGCTCGGGCTGTG	900
DB	281	HisGlySerLeuValAsnSerTrpThrAsnThrProTrpThrGlyAlaLeuGlyLeuLeu	300

RESULT 13

QY	901	GACTTTGCGCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGTC	960
DB	301	AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal	320
QY	961	TCCGTTATTCCAGCACTCTCGCCACCGCTTGTGCGGGTCGCGAGCGGATGCGGAG	1020
DB	321	SerArgTrpSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu	340
QY	1021	CTCACCACACGGCTGCTACCGCTTATGTAAGGACCTCTATTATTACTAGTACTATGGT	1080
DB	341	LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTrpPheThrSerThrAsnGly	360
QY	1081	GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTTCTGACACTCTGCTT	1140
DB	361	ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu	380
QY	1141	GGCGGCTCCGACAGAAATGTATTCGTGCGTGGTGGCGAGCTGTTACTCCCGTCCC	1200
DB	381	GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTrpSerArgPro	400
QY	1201	GTGTCTCAGCAATGGCGGCGGCTGTAAGTTGTATACATCTCTAGAGATGCTCAG	1260
DB	401	ValValSerAlaAsnGlyGluProThrValLysLeuTrpThrSerValGluAsnAlaGln	420
QY	1261	CAGGATAAGGCTATTGCAATCCGATGACATTGACCTCGAGAAATCTCGTGTGGTTAT	1320
DB	421	GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle	440
QY	1321	CAGGATTATGATAACCAACATGAACATCGCGCGACGCTTCTCCAGCCCATCGCGC	1380
DB	441	GlnAspTrpAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg	460
QY	1381	CCTTCTCTCTCTCGAGCTAATGATGCTTGTGCTCTCTCTACCGCTGCGGATAT	1440
DB	461	ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTrp	480
QY	1441	GCACGCTCCACTTATGGCTCTTCGACTGGCCAGATTATGTTTCTGACTCTGTGACCTTG	1500
DB	481	AspGlnSerThrTrpGlySerSerThrGlyProValTrpValSerAspSerValThrLeu	500
QY	1501	GTTAATGTTGCGACCGCGCGCGCGCTGTCGCTCGATTGGACCAAGGTCACA	1560
DB	501	ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr	520
QY	1561	CTTGAGGTCGCGCCCTCTCCACCATCCAGCAGTACTCGAGACCTCTTGTGCTGCGG	1620
DB	521	LeuAspGlyArgProLeuSerThrIleGlnGlnTrpSerLysThrPhePheValLeuPro	540
QY	1621	CTCCGCGGTAAGCTCTTCTTCTGGGAGGCGAGCACAACTAAAGCCGGGTACCCCTTAAAT	1680
DB	541	LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTrpProTrpAsn	560
QY	1681	TATAACACCACTGCTAGCAACCAACTGCTTGTGCGAATGCGCGCGCGGACCGGTCGCT	1740
DB	561	TrpAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla	580
QY	1741	ATTTCACACTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1800
DB	581	IleSerThrTrpThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal	600
QY	1801	TTAGCCCCCACTCTGCGCTAGCATTTGTTAGGATACCTTCGACTTACCTGCGCGCGC	1860
DB	601	LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTrpProAlaArgAla	620
QY	1861	CATACCTTTTGTATTTCTGCCAGAGTCCGCCCTTTGGCTTTCAGGGCTGCGCTTTC	1920
DB	621	HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe	640
QY	1921	CAGTCTACTGTCGTGAGCTTCAGCGCTTACAGTGAAGGTCGCTAAACTCGGAGTTG	1980
DB	641	GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu	660

Db 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuVal 460
QY 1714 GAGATGCCCGCGGACCGGTCGCTATTTCACCTTACACCACTAGCCTGGTCTGT 1773
Db 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGly 480
QY 1774 CCGCTCTCCATTCTCGCGTTCGCGTTAGCCGCCCACTCTCGCTAGCATTTGCTTGA 1833
Db 481 ProValSerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuLeuGlu 500
QY 1834 GATACCTTGGACTACCTGCCCGGCCCATCTTTTGATGATTTCTGCCAGAGTCCGC 1893
Db 501 AspThrLeuAspTyrProAlaArgAlaHisThrPheAspPheCysProGluCysArg 520
QY 1894 CCGCTTGGCTTCAGGCTCGGCTTCCAGTCTACTGTGCTGAGCTTCAGCGCCTTAAG 1953
Db 521 ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGlnLeuGlnArgLeuLys 540
QY 1954 ATGAGGTGGTAAACTCGGAGTTG 1980
Db 541 MetLysValGlyLysThrArgGluLeu 549

RESULT 14

US-08-477-292-15
; Sequence 15, Application US/08477292
; Patent No. 6291641
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas
; APPLICANT: McAttee, Patrick
; APPLICANT: Yarbough, Patrice
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cathleen M. Desjardins, M.D.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: US
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,292
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/327,952
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, Allan A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: G32P5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)369-9500
; TELEFAX: (415)368-0709
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 549 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
; INDIVIDUAL ISOLATE: r62kDa, FIGURE 4
US-08-477-292-15
Alignment Scores:

Pred. No.: 0 Length: 549
Score: 549.00 Matches: 549
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.18% Indels: 0
DB: 3 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-08-477-292-15 (1-549)
QY 334 GGGTGGTCCGGCCCATGACACCCCGCAGTGCCTGATGTGACTCCCGCGCGCATC 393
Db 1 AlaValAlaProAlaHisAspThrProValProAspValAspSerArgGlyAlaIle 20
QY 394 TTGCGCGGCACTATAACCTATCAACATCTCCCTTACCTCTTCGTTGCCACCGCACT 453
Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThr 40
QY 454 AACCTGGTTCTTTATGCGCCCGCTCTTATGTCGCGCTTTTACCCCTTCAGACGCGCAAT 513
Db 41 AsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACCATATAATGCGCACGGAAGCTTCTAATTATGCCAGTACCGGTTGCCGTGCCACA 573
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThr 80
QY 574 ATCCGTTACCGCGCTGCTGCCCAATGCTGCGGGTTAGCCCATCTCCATCTCATTC 633
Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCCACAGACACACACACCCCGAGCTCCGTTGATATGAATTCATAACCTCGACGAT 693
Db 101 TrpProGlnThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120
QY 694 GTTGTATTTTATGTCAGCCCGGCTAGCTCTGAGCTTGTGATCCCAAGTAGCGCCTA 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTATCGTAACCAAGGCTGCGCTCCGTCGAGACCTCTGGGTGCTGAGGAGGAGCT 813
Db 141 HisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCTGGTCTTTGTTATGCTTTGCATACATGGCTCAGTAAATTCCTATATATA 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180
QY 874 CCGTATACCGGTGCCCTCGGCTGTGGACTTTGGCTTGGCTTGGCTTGGCTTGGCTT 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200
QY 934 ACCCGGTTAACCAATACCGGGTCTCCGTTATTTCCAGACCTCTCGCCACCGCCTT 993
Db 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerThrThrAlaArgHisArgLeu 220
QY 994 CGTGGGTGCGGACGGGACTCCGAGCTCACCACCGGCTGCTACCGCTTTTATGAAG 1053
Db 221 ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLys 240
QY 1054 GACCTCTATTTACTAGTACTAATGGTGTGAGATCGGCGCGGATAGCCCTCACC 1113
Db 241 AspLeuTyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThr 260
QY 1114 CTGTTCAACCTTGTGACACTCTGCTTGGCGGCTCGCCAGACAGAAATTGTTGCTGGCT 1173
Db 261 LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuLeuSerSerAla 280
QY 1174 GTGGCCAGCTGTTCTTACTCCCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1233
Db 281 GlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLys 300
QY 1234 TTGTATACATCTGTAGAGAAATGCTCAGCAGGATAGGGTATTGCAATCCCGCATGACAT 1293
Db 301 LeuTyrThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIle 320

```
QY 1294 GACCTCGAGAAATCTCTGTGGTTATTACAGATTATGATACCAACATGAACAGATCGG 1353
Db 321 AspLeuGlyGluSerArgValValleGlnAspTyrAspGlnHisGluGlnAspArg 340
QY 1354 CCGAGCGCTTCTCCAGCCCATCGCGCCCTTCTCTGTCTCCAGCTAATGATGCTT 1413
Db 341 ProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu 360
QY 1414 TGGCTCTCTCACCCTGCGGAGTAGACAGTGCACCTATGCTCTTCAGTGGCCCA 1473
Db 361 TrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPro 380
QY 1474 GTTTATCTTCTGACTCTGTGACCTCTGTTAAATGTTGCGCGCGCAGCGCTGGCC 1533
Db 381 ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400
QY 1534 CGGTCTCTGATGGACCAAGGTGCACACTTGCAGCGTGCCTCCCTCTCCACCATCCAGCAG 1593
Db 401 ArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420
QY 1594 TACTCGAGACCTCTTGTCTGCGCTCGCGGTAAAGCTCTCTTCTGGAGGAGGC 1653
Db 421 TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGly 440
QY 1654 ACACTAAAGCGGGTACCCCTTATAATTATAACACCACTGCTAGCGACCAACTGCTTGTG 1713
Db 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuVal 460
QY 1714 GAGAAATCGCGCGGACCGGGTCCGCTATTTCACCTTACACCACTACGCTGGGTGGT 1773
Db 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGly 480
QY 1774 CCGCTCTCCATTCTCGCGTTCGCGTTTACGCCCTAGCCCTGCGTAGCATTCCTTGAG 1833
Db 481 ProValSerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGlu 500
QY 1834 GATACCTTGACCTACCCCTGCGCGCCCATCTTTCATGATTTCTGCCAGAGTGCCTG 1893
Db 501 AspThrLeuAspTyrProAlaArgAlaHisThrPheAspAspPheCysProGluCysArg 520
QY 1894 CCCCTTGGCTTCAGGCTCGCTTTCAGTCTACTGCTGCTAGCTTACGCGCTTAAG 1953
Db 521 ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLys 540
QY 1954 ATGAAGTGGGTAAACTCGGGAGTTG 1980
Db 541 MetLysValGlyLysThrArgGluLeu 549
```

RESULT 15

PCT-US95-13703-15

; Sequence 15, Application PC/TUS9513703

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND

; TITLE OF INVENTION: USES THEREFOR

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates

; STREET: P.O. Box 60850

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306-0850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/13703

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

```
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 549 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
; INDIVIDUAL ISOLATE: r62kDa, FIGURE 4
; PCT-US95-13703-15
```

Alignment Scores:

Pred. No.:	0	Length:	549
Score:	549.00	Matches:	549
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	83.18%	Indels:	0
DB:	5	Gaps:	0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x PCT-US95-13703-15 (1-549)

QY	334	CGCGTCTCGCTCCGCGCCATGACACCCCGCAGTGCCTGATGCGACTCCCGCGGCGCATC	393
Db	1	AlaValAlaProAlaHisAspThrProValProAspValAspSerArgGlyAlaIle	20
QY	394	TTGCGCGGCGAGTATAACCTATCAACATCTCCCTTACCTCTCCGTGGCCACCGGCACT	453
Db	21	LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerValAlaThrGlyThr	40
QY	454	AACCTGTTCTTATGCGCGCCCTCTTATGTCGGCTTTTACCCCTTACAGGACGGCACCAAT	513
Db	41	AsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn	60
QY	514	ACCCATATAATGCGCAGGAGCTTCTTAATTATGCCAGTACCGGGTTCGCCGTGCCACA	573
Db	61	ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThr	80
QY	574	ATCGGTTACCGCGCTGGTCCCCCATGCTGTGCGGGGTTACGCCATCTCCATCTCATTC	633
Db	81	IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe	100
QY	634	TGGCCACAGACACACCCACCGACGTCGTTGATATGAATCAATAACTCGACCGAT	693
Db	101	TrpProGlnThrThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp	120
QY	694	GTTCTGATTTTACTGCAGCCCGCATAGCTCTGAGCTTGTGATCCCAAGTAGCGCCTA	753
Db	121	ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu	140
QY	754	CACATATCGTAACCAAGGCTCGCTCCGTGAGACCTCTGGGTGGTGGTGGAGGAGGCT	813
Db	141	HisTyrArgAsnGlnGlyTrpArgSerValGluThrSerGlyValAlaGluGluAla	160
QY	814	ACCTCTGGTCTTGTATGCTTTTCATACATGGCTCACTCGTAATTCCTATCTACTATACA	873
Db	161	ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerTyrThrAsnThr	180
QY	874	CCCTATACCGGTGCGCTCGGGCTGTGGACTTTGCCCTTACGCTTGAGTTTCGCAACCTT	933
Db	181	ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu	200
QY	934	ACCCCGGTAAACACCAATACCGGGTCTCCGTTTATTCAGCACTGCTCCACCGGCTT	993
Db	201	ThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu	220

994 CGTCGGGTGGGACGCGAGCTCCAGCTCACCACCGGTGCTACCGCTTTATGAAG 1053
|||
Db ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLys 240
1054 GACCTCTATTTTACTAGTAAATGGTCTCGGTGAGATCGGCGCGGATAGCCCTCACC 1113
|||
Db AspLeuTyrPheThrSerThrAsnGlyValGlyGluLeuGlyAlaLeuThr 260
1114 CTGTTCAACCTTGTGACACTCTGCTGGCGGCTCCGACAGAAATGATTTCGCGCT 1173
|||
Db LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuLeuSerSerAla 280
1174 GGTGCCAGCTGTTCTACTCCGTCCTCCGTTGCTCAGCAATGGCGAGCGACTGTTAAG 1233
|||
Db GlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLys 300
1234 TTGTATACATCTGTAGAGAATGCTCAGCAGGATAAGGGTATTGCAATCCGCGATGACATT 1293
|||
Db LeuTyrThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIle 320
1294 GACCTCGAGAAATCTGCTGTGTTATTCAGGATTATGATAACCAACATGAACAAGATCGG 1353
|||
Db AspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArg 340
1354 CCGACGCTTCTCCAGCCCATCGCGCCCTTCTCTGCTCCTTCGAGCTAATGATGTGCTT 1413
|||
Db ProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeu 360
1414 TGGCTCTCTCTCACCGTCCGAGATGACACAGTCCACTTATGGTCTTTCGACTGGCCCA 1473
|||
Db TrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerThrGlyPro 380
1474 GTTATGTTTCTGACTCTGTGACCTTGGTATATCTTGGACCGCGCGGCGGCTTGC 1533
|||
Db ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400
1534 CGGTCCGCTCGATTGACCAAGGTCACTTGACGCTCGCCCTCTCCACCATCCAGCAG 1593
|||
Db ArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420
1594 TACTCGAAGACCTCTTTGCTCGCTCCGCTCCGCGTAAAGTCTCTTTCTGGAGGCGGC 1653
|||
Db TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGly 440
1654 ACACTAAAGCGGTACCTTATATTAACACCACTGCTAGCACCACCACTGCTTCTC 1713
|||
Db ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuVal 460
1714 GAGATGCCCGCGGCACCGGTGCTATTTCACCTTACACCACTAGCCTGGGTGCTGGT 1773
|||
Db GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGly 480
1774 CCCGTCCTCCATTCTCGGTTCCGTTTTCAGTTCGCTTTCAGTTCGCTTTCAGTTCGCT 1833
|||
Db ProValSerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuGlu 500
1834 GATACCTGGACTACCTCCCGCGCCCATCTTTTGATGATTCTGCGCAGAGTCCCGC 1893
|||
Db AspThrLeuAspTyrProAlaArgAlaHisThrPheAspPheCysProGluCysArg 520
1894 CCCCTGGGCTTCAGGCTCGCTTTCAGTTCGCTTTCAGTTCGCTTTCAGTTCGCTTTCAG 1953
|||
Db ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLys 540
1954 ATGAGGTGGTAAACTCGGAGTTG 1980
|||
Db MetLysValGlyLysThrArgGluLeu 549

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:55:54 ; Search time 97.3236 Seconds
(without alignments)
12792.733 Million cell updates/sec

Title: US-09-851-410a-6_COPY_5147_7129
Perfect score: 660
Sequence: 1 ATGCGCCCTCGGCTATT.....GTAAACTCGGAGTTGTAG 1983

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1292805 seqs, 313927144 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2482574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09851410/runat_16082004_125719_9771/app_query.fasta_1.4238
-DB=Published Applications AA -QPMF=fastan -SUFFIX=oli.rapb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40 cdi -LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09851410@cgn2_1_1_27@runat_16082004_125719_9771
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PTC_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PTC_US_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
---------------	-------	----------------	--------	----	-------------

1	660	100.0	660	9	US-09-769-066-13	Sequence 13, Appl
2	660	100.0	660	10	US-09-851-410-8	Sequence 8, Appl
3	660	100.0	660	14	US-10-185-868-19	Sequence 19, Appl
4	549	83.2	549	9	US-09-769-066-15	Sequence 15, Appl
5	541	82.0	660	15	US-10-257-044-1	Sequence 1, Appl
6	540	81.8	540	9	US-09-769-066-25	Sequence 25, Appl
7	525	79.5	525	9	US-09-769-066-27	Sequence 27, Appl
8	436	66.1	436	14	US-10-165-868-17	Sequence 17, Appl
9	408	61.8	660	12	US-10-381-770-1	Sequence 1, Appl
10	327	49.5	327	9	US-09-769-066-17	Sequence 17, Appl
11	327	49.5	327	14	US-10-165-868-15	Sequence 15, Appl
12	210	31.8	214	12	US-10-381-770-2	Sequence 2, Appl
13	210	31.8	232	12	US-10-381-770-12	Sequence 12, Appl
14	200	30.3	389	12	US-10-381-770-4	Sequence 4, Appl
15	190	28.8	194	12	US-10-381-770-3	Sequence 3, Appl
16	165	25.0	169	12	US-10-381-770-7	Sequence 7, Appl
17	155	23.5	159	12	US-10-381-770-8	Sequence 8, Appl
18	145	22.0	149	12	US-10-381-770-9	Sequence 9, Appl
19	135	20.5	139	12	US-10-381-770-10	Sequence 10, Appl
20	120	18.2	525	9	US-09-769-066-28	Sequence 28, Appl
21	120	18.2	540	9	US-09-769-066-26	Sequence 26, Appl
22	120	18.2	549	9	US-09-769-066-16	Sequence 16, Appl
23	120	18.2	659	14	US-10-165-868-20	Sequence 20, Appl
24	120	18.2	660	9	US-09-769-066-14	Sequence 14, Appl
25	118	17.9	674	16	US-10-239-090A-55	Sequence 55, Appl
26	118	17.9	674	16	US-10-239-090A-57	Sequence 57, Appl
27	109	16.5	123	10	US-09-851-410-9	Sequence 9, Appl
28	109	16.5	123	12	US-10-381-770-11	Sequence 11, Appl
29	109	16.5	124	9	US-09-769-066-21	Sequence 21, Appl
30	78	11.8	123	14	US-10-224-999A-3478	Sequence 3478, Ap
31	75	11.4	660	10	US-09-468-147-92	Sequence 92, Appl
32	75	11.4	660	12	US-10-319-745-92	Sequence 92, Appl
33	75	11.4	660	16	US-10-239-090A-51	Sequence 51, Appl
34	75	11.4	660	16	US-10-239-090A-53	Sequence 53, Appl
35	61	9.2	124	9	US-09-769-066-22	Sequence 22, Appl
36	61	9.2	327	10	US-09-468-147-175	Sequence 175, App
37	61	9.2	327	10	US-09-468-147-176	Sequence 176, App
38	61	9.2	327	12	US-10-319-745-175	Sequence 175, App
39	61	9.2	327	12	US-10-319-745-176	Sequence 176, App
40	61	9.2	338	10	US-09-468-147-199	Sequence 199, App
41	61	9.2	338	10	US-09-468-147-200	Sequence 200, App
42	61	9.2	338	12	US-10-319-745-199	Sequence 199, App
43	61	9.2	338	12	US-10-319-745-200	Sequence 200, App
44	61	9.2	459	10	US-09-468-147-206	Sequence 206, App
45	61	9.2	459	10	US-09-468-147-207	Sequence 207, App

ALIGNMENTS

RESULT 1
US-09-769-066-13
; Sequence 13, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; Mcatee, C. Patrick
; Yarbrough, Patrice O.
; Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

Db 541 LeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATACACACCTAGCGACCAACTGCTTTCGAGATCCCGCGGACCGGTCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaGlyHisArgValAla 580
QY 1741 ATTTCACACTACACACTAGCCTGGGTGCTGCTCCGCTCCATTCTCGGTTGCCGTT 1800
Db 581 IleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTGCGCTAGCATGCTTGGATACCTTGGACTACCTCCCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACCTTTGATGATTTCTGCCAGAGTCCGCCCTTGGCTTCAGGGCTGCGCTTTC 1920
Db 621 HisThrPheAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGAGCTTCAGCGCCTTAAGATCAAGGTGGTAAACTCGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660
RESULT 2
US-09-851-410-8
; Sequence 8, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yarbrough, Patrice O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; Fty, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,410
; FILING DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,898
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-851-410-8
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 660.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-851-410-8 (1-660)
QY 1 ATGCGCCCTCGGCGCTATTTTGTGCTGCTCTCATGTTTTCCTATGCTGCGCGCCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCGCCGCTCAGCCGCTCTGGCGCGCTGCTGGCGCGCGCAGCGCGGTTCGGCGGTGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGlyGly 40
QY 121 TTCTGGGTGACCGGGTGTATCTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCGCCCGATGTCACCGCTCGCGCGGGGTGACCTCGTGTTCGCAACCGCGC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCTTGGCGTGACGAGGCCCGCGCGCGCGCTTCACGTGCT 300
Db 81 ArgProLeuGlySerAlaTyrArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACTTACACAGCTGGGGCGCGCGCTAACCGCGGTGCTCGCGCGCGCATGACACCGG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTCCGTGATGTCGACTCCGCGCGCGCCATCTTGGCGCGCGCAGTATACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACTCTCTCTCGTGGCCACCGCGCACTAACCTGGTTCCTTTATGCGCGCCCTCT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCGTTTACCCCTTCAGGCGGACCAATACCATATAATGCGCGCAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATATGCCAGTACCGGGTTGCGCGTCCCAATCCGTTACCGCGCGCTGGTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTGCGCGGTACGCCATCTCCATCTCATCTCTGGCCACAGACACCAACCGCGCGG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATGATTCATTAACCTCAGCGGATGCTGCTATTTAGTCAGCGCGCGGATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCGAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTCAGGAGGAGGCTACCTGCTGTTGTTATGCTTTGCATA 840

Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
 QY 841 CATGGCTCACTGTAATTCCTATACCTATACACCTATACCGGTGCCCTCGGGCTGTG 900
 Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
 QY 901 GACTTTGGCTTCAGCTTCAGTTTCGCACTTACCCCGGTAAACCAATACCGGGTC 960
 Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
 QY 961 TCCCGTTATTCCAGCACTGCTCCGCCACCGCTTCGTCGGGTCCGGACGGAGTCCGAG 1020
 Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
 QY 1021 CTACACACACGGCTGCTACCCGCTTTATGAAGAGACTCTATTTTACTAGTACTATGGT 1080
 Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
 QY 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTTGTCTGACACTCTGCTT 1140
 Db 361 ValGlyGluLeuGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
 QY 1141 GCGCGCTCCGACAGAAATGATTTCGTGGGTGGCGAGCTGTCTACTCCCGTCC 1200
 Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyLeuLeuPheTyrSerArgPro 400
 QY 1201 GTTGTCTCAGCAATCGCGAGCGACTGTGTTAAGTTGTATACATCTGTAGATGCTCAG 1260
 Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
 QY 1261 CAGGATAAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAACTCTGTGTGTTATT 1320
 Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 440
 QY 1321 CAGGATTATGATAACCAACATGAACAGATCGCGACGCTTCTCCAGCCCATCGCGC 1380
 Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
 QY 1381 CCTTCTCTGCTCTCGAGTATGATGCTGTTGGCTCTCTCCACCGGTCCGGAGTAT 1440
 Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
 QY 1441 GACCACTGCTTATGCTCTCGACTGGCCCACTTATGTTTCTGACTCTGTGACCTTG 1500
 Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
 QY 1501 GTTAATGTCGACCGCGCGCGAGCGTTCGCGGTGCGTTCGATGACACAGGTCACA 1560
 Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
 QY 1561 CTTGACGGTCCGCTCTCCACCATCCAGCACTCTCGAAGACCTCTTTGTCCTGCGCG 1620
 Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
 QY 1621 CTCGGGGTAAGCTCTCTTTCTGGGAGCGGACCACTAAAGCCGGGTACCTCTATAAT 1680
 Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
 QY 1681 TATAACACCACTGCTAGCGACCACTGCTTGTGAGAAATCGCGCGGACCGCGGTGCT 1740
 Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
 QY 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCGCTCTCCATTTCTGGGTTCGCGTT 1800
 Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
 QY 1801 TTAGCCCCCACTCTGCGCTAGCATTCCTTGGAGATACCTTGGACTACCTCGCCGCGCC 1860
 Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
 QY 1861 CATACTTTGATGATTTCTGCCAGAGTGGCGGCTTGGCTTCCAGGCTCGGCTTTC 1920
 Db 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640

QY 1921 CAGTCTACTGTCCTCAGCTTCAGCGCCTTAAGATGAAGTGGGTAAAACTCGGAGTTG 1980
 Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetCysValGlyLysThrArgGluLeu 660
 RESULT 3
 US-10-165-868-19
 : Sequence 19, Application US/10165868
 : Publication No. US20030143241A1
 : GENERAL INFORMATION:
 : APPLICANT: Reyes, Gregory R.
 : Bradley, Daniel W.
 : Twu, Jr-Shin
 : Purdy, Michael A.
 : Tam, Albert W.
 : Krawczynski, Krzysztof Z.
 : TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
 : NUMBER OF SEQUENCES: 22
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Dehlinger & Associates
 : STREET: 350 Cambridge Avenue, Suite 250
 : CITY: Palo Alto
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 94306
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/10/165,868
 : FILING DATE: 06-Jun-2002
 : CLASSIFICATION: <Unknown>
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/07/870,985A
 : FILING DATE: 20-APRIL-1992
 : APPLICATION NUMBER: US 822,335
 : FILING DATE: 17-JAN-1992
 : APPLICATION NUMBER: US 505,888
 : FILING DATE: 05-APRIL-1990
 : APPLICATION NUMBER: US 420,921
 : FILING DATE: 13-OCTOBER-1989
 : APPLICATION NUMBER: US 367,486
 : FILING DATE: 16-JUNE-1989
 : APPLICATION NUMBER: US 336,672
 : FILING DATE: 11-APRIL-1989
 : APPLICATION NUMBER: US 208,997
 : FILING DATE: 17-JUNE-1988
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Fabian, Gary R.
 : REGISTRATION NUMBER: 33,875
 : REFERENCE/DOCKET NUMBER: 4600-0093.30
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 324-0880
 : TELEFAX: (415) 324-0960
 : INFORMATION FOR SEQ ID NO: 19:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 660 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : HYPOTHETICAL: NO
 : ORIGINAL SOURCE:
 : INDIVIDUAL ISOLATE: ORF 2, BURMA, FIGURE 9
 : SEQUENCE DESCRIPTION: SEQ ID NO: 19:
 : US-10-165-868-19
 Alignment Scores:
 Pred. No.: 0 Length: 660
 Score: 660.00 Matches: 660
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-165-868-19 (1-660)

QY 1 ATGGCGCCCTCGCCGCTATTTGTTGCTGCTCCTCATGTTTTCCTATGCTGCGCGGCCA 60
DB |||||
QY 1 MetArgProArgProlleuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
DB |||||
QY 61 CGGCGCGGTACGCGTCTGCGCGCGCGTCTGCGCGCGCGAGCGCGGTTCGCGCGGTGT 120
DB |||||
QY 21 ProProGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlySerGlyGlyGly 40
DB |||||
QY 121 TTCTGGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
DB |||||
QY 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
DB |||||
QY 181 CCTTTGCGCCCGATGTACACGCTCGCGCGCGGCTGGACCTCTGTTTCGCGCAACCGCC 240
DB |||||
QY 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
DB |||||
QY 241 CGACACTCGGCTCGGTTGGGTGACACAGGCCCGAGCGCGCGCGCTGCTGCTCAGTGT 300
DB |||||
QY 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
DB |||||
QY 301 AGACCTACACAGCTGGCGCGCGCGCTAACCGCGGTCTCGCGCGCGCATGACACCGCG 360
DB |||||
QY 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
DB |||||
QY 361 CCAGTGCCTGATGTGCACTCCCGCGCGCGCATCTTTGCGCGCGCAGTATACCACTCAACA 420
DB |||||
QY 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
DB |||||
QY 421 TCTCCCTTACTCTTCGCGTGGCCACCGGCACTAACCTGGTCTTATTATGCGCGCGCTCT 480
DB |||||
QY 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
DB |||||
QY 481 AGTCGGCTTTTACCCTTCAGAGCGCACCATACCATATATATGCGCACGGAAGCTCT 540
DB |||||
QY 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
DB |||||
QY 541 AATTATGCCCACTACCGGCTGCGCGTGCACAAATCCGTTACCGCGCGGTGCTCCCAAT 600
DB |||||
QY 181 AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValProAsn 200
DB |||||
QY 601 GCTGTCGGGGTTACGCCATCTCCATCTCTCTGCGCACAGACACACACCGCGAGC 660
DB |||||
QY 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
DB |||||
QY 661 TCCGTTGATATCAATTCAATACCTCGAGCGATGTTCTGTTATTTAGTCCAGCGCGCAT 720
DB |||||
QY 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
DB |||||
QY 721 GCCTCTGAGCTTGTATCCCAAGTAGCGCCCTACACTATCGTAACCAAGCTGCGCGCTCC 780
DB |||||
QY 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
DB |||||
QY 781 GTCGAGACTCTGGGGTGGCTGAGGAGGAGGCTACTCTGCTGCTTGTATTGCTTGCATA 840
DB |||||
QY 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
DB |||||
QY 841 CATGCTCACTCGTAATTCCTATATAACCTATACACCTATACCGGTGCTCGGCTGTTG 900
DB |||||
QY 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
DB |||||
QY 901 GACTTTGCGCTTGAGCTTGAGTTTCGCAACCTTACCGCGGTAAACACCAATACGGGGTTC 960
DB |||||
QY 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
DB |||||
QY 961 TCCGTTATTCCAGCACTGCTGCCACCGCTTGTGCGGGTGGCGAGCGGACTGCGCGAG 1020
DB |||||
QY 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
DB |||||

1021 CTCACCAACGAGGTGCTACCGGCTTTATGAGGACCTCTATTTTACTAGTACTAATGTT 1080
DB |||||
QY 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
DB |||||
QY 1081 GTCGGTGAATGCGCGCGGATAGCCCTCACCCTGTTTCAACCTTGCTGACACTCTGCTT 1140
DB |||||
QY 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
DB |||||
QY 1141 GCGCGCTGCGCAGCAGCAATTCATTTCGTCGCGTGGTGGCCAGCTGCTTACTTCCCGTCCC 1200
DB |||||
QY 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlnLeuPheTyrSerArgPro 400
DB |||||
QY 1201 GTTGTCTCAGCAATGCGCAGCCAGCTGTTAAGTTGTATACATCTGTAGAGAACTCTCAG 1260
DB |||||
QY 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
DB |||||
QY 1261 CAGGATAAGGTATTGCAATCCCGCATGACATTGACCTCGGAGAACTCTGTTGTTATT 1320
DB |||||
QY 421 GlnAspLysGlyIleAlaIleProHisAspLeuGlyGluSerArgValValIle 440
DB |||||
QY 1321 CAGGNTATTGATAACCAACATGAAACATGCGCGCGCTTCTCCAGCCCGCATCGCGC 1380
DB |||||
QY 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
DB |||||
QY 1381 CCTTTCTCTGCTCCTTCGAGCTAATGATGCTGCTTTGGCTCTCTCTCACCGCTGCCAGTAT 1440
DB |||||
QY 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
DB |||||
QY 1441 GACAGTCCACTTATGCTCTTCGACTGCGCGCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
DB |||||
QY 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
DB |||||
QY 1501 GTTAAATGTTGCGCGCGCGCGCGCTGCGCGCTGCTGCGTTCGATTGGACCAAGTGCACA 1560
DB |||||
QY 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
DB |||||
QY 1561 CTTGACGCTGCGCGCGCTCTCCACCATCCAGCAGTACTCGAAGACCTCTCTTGTCTCGCG 1620
DB |||||
QY 521 LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro 540
DB |||||
QY 1621 CTCGCGGTAGCTCTCTTTCTGGAGGAGGAGCACAACCTAAAGCGGGTACCTTATAT 1680
DB |||||
QY 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
DB |||||
QY 1681 TATAACACACTCTGAGCGACCACTGCTGTCGAGAAATGCGCGCGGACACCGGTCGCT 1740
DB |||||
QY 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
DB |||||
QY 1741 ATTTCCACTTACACCATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB |||||
QY 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
DB |||||
QY 1801 TTAGCCCCCACTCTGCGCTAGCATGCTTGGAGTATACCTTGGACTACCTGCGCGCGC 1860
DB |||||
QY 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
DB |||||
QY 1861 CATATCTTTGATGATTTCTGCGCAGTGCGCGCGCTTGGCTTTCAGGGCTGCGCTTTC 1920
DB |||||
QY 621 HisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
DB |||||
QY 1921 CAGTCTACTGCTGCTGAGCTTACAGCTTAAAGTGAAGGTGGGTAAACTCGGAGTTG 1980
DB |||||
QY 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660
DB |||||

RESULT 4
US-09-769-066-15
; Sequence 15, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; McAtee, C. Patrick
; Yarbough, Patrice O.
; Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```
QY 1774 CCGGCTCCATTTCTGCGGTTGCGCTTTAGCCCGCCACTCTGCGCTAGCATTGCTTGAG 1833
DB 481 ProValSerIleSerAlaValAlaValIleuAlaProHisSerAlaLeuAlaLeuLeuGlu 500
QY 1834 GATACCTTGGACTACCCCTCGCGCCGCCATACCTTTTGATGATTTCTGCGCCAGAGTGGCGC 1893
DB 501 AspThrLeuAspTyrProAlaAlaAlaHisThrPheAspAspPheCysProGluCysArg 520
QY 1894 CCCCTTGGCTTCAGGGCTGCGCTTTCAGTCTACTGCTGCTGAGCTTCAGCGCTTAAG 1953
DB 521 ProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLys 540
QY 1954 ATGAAGTGGTAAACCTCGGAGTTG 1980
DB 541 MetLysValGlyLysThrArgGluLeu 549

RESULT 5
US-10-257-044-1
; Sequence 1, Application US/10257044
; Publication No. US20030220475A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; APPLICANT: Represented by the Secretary, Department of Health and
; APPLICANT: Human Services, Centers for Disease Control and Prevention
; APPLICANT: HOWARD A. FIELDS
; APPLICANT: YURI E. KHUDYAKOV
; APPLICANT: JIHONG MENG
; TITLE OF INVENTION: NEUTRALIZING IMMUNOGENIC HEV POLYPEPTIDES (HEV)
; FILE REFERENCE: 14114.034002
; CURRENT APPLICATION NUMBER: US/10/257,044
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,380
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 660
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: / No. US20030220475A1e =
US-10-257-044-1
Alignment Scores:
Pred. No.: 0 Length: 660
Score: 541.00 Matches: 641
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 81.97% Indels: 0
DB: 15 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-257-044-1 (1-660)
QY 55 GCGCACCGCGGTCAGCGCTGCGCGCGCTGCTGCGCGCGCGCGCGCGGTTCCGCGC 114
DB 19 AlaProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGly 38
QY 115 GGTGTTTCTGGGTGACCGGGTTGATCTCAGCCCTTCGCAATCCCTTATATTCATCCA 174
DB 39 GlyGlyPheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisPro 58
QY 175 ACCAACCCCTTCGCGCGCGATGTCACCGTGGCGCGCGGCTGAGACCTGTTTCGCGAA 234
DB 59 ThrAsnProPheAlaProAspValThrAlaAlaGlyAlaGlyProArgValArgGln 78
QY 235 CCGCGCGGACCACTCGGCTCGCTTGGCTGACGAGGCCCGCGCGCGCTTCCTCA 294
DB 79 ProAlaArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAsp 98
QY 295 CGTCTAGACCTACACAGCTGGGCGCGCGCTAAACCGCGCTCGCTCCGCGCCATGAC 354
```

```
DB 99 ArgArgArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAsp 118
QY 355 ACCCGCGCAGTGCCTGATGTGACTCCGCGCGCGCATCTTTCGCGCGGAGTATACCTA 414
DB 119 ThrProProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeu 138
QY 415 TCAACATCTCCCTTACCTCTTCCTGCGCGCACCGGCACTAACTGCTGTTCTTTATCGCGC 474
DB 139 SerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAla 158
QY 475 CTTCTTACTGCGCTTTTACCCCTTCAGACGCGCAACCAATACCATATATATGCGCCAGAA 534
DB 159 ProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGlu 178
QY 535 GCTTCTAATTATGCCCCAGTACCGGTTGCGCGTGCCACAATCCGTTACCGCGCGTGTGC 594
DB 179 AlaSerAsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuVal 198
QY 595 CCCAATGCTGTCGCGGGTTAGCCCATCTCCATCTCATTTCTGGCCACAGACACCAACCACC 654
DB 199 ProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThr 218
QY 655 CCGACGCTCGTGTGATGATGAATTCATAACCTCGACGAGTGTTCGTAATTTAGTCACGCC 714
DB 219 ProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnPro 238
QY 715 GGCATAGCTCTGAGCTTGTGATCCCAAGTAGCGCCCTACACTATCTGTAACCAAGCTGG 774
DB 239 GlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrp 258
QY 775 CGCTCCGTCGAGACCTCGGGTGGCTGAGGAGGAGGTACTCTGCTGCTTCTTTATGCTT 834
DB 259 ArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeu 278
QY 835 TGCATACATGGCTCACTCGTAAATTCCTATACATAACACCCCTATACCGGTCCTCGGG 894
DB 279 CysIleHisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGly 298
QY 895 CTGTTGGACTTTGCGCTTGAGCTTGAGTTGCAACCTTACCCCGGTTAACCAACATAGC 954
DB 299 LeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThr 318
QY 955 CGGCTCTCCGCTTATTCGAGCACTCTCTCGCACCGCTTCTGTCGCGGTGCGACGGACT 1014
DB 319 ArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThr 338
QY 1015 GCGAGCTCACACACCGCTGCTACCGCTTATGAAGGACCTCTATTTTACTAGTACT 1074
DB 339 AlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThr 358
QY 1075 AATGCTGTCGCTGAGATCGCGCGGGATAGCCCTCACCCCTGTTCAACCTTGTGACACT 1134
DB 359 AsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThr 378
QY 1135 CTGCTTGGCGGCTCGCCACAGAAATTTGATTTCTGCTGCTGGTGGCAGCTGTTCTACTCC 1194
DB 379 LeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyCysGlnLeuPheTyrSer 398
QY 1195 CGTCCCGTTGTCTCAGCCAAATGGCGAGCGCGACTGTTAAAGTTGTATACATCTGTAGAGAT 1254
DB 399 ArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsn 418
QY 1255 GCTCAGGAGGATAGGGTATTGGCAATCCCGCATGACATTGACCTCGGAGAAATCTGCTGT 1314
DB 419 AlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgVal 438
QY 1315 GTTATTACAGGATATGATAACCAACATGAACAAGATCGCGCGCGCTTCTCCAGCCCA 1374
DB 439 ValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaPro 458
QY 1375 TCGCGCCCTTTCTCTGTCTCTCGAGCTAATGATGTGCTTTGGCTCTCTCTCTCACCGCTGCC 1434
DB 459 SerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAla 478
```

1435 GAGTATGACGAGTCCACTTATGGCTCTTCGACTGGCCCGAGTTTATGTTCTGACTCTG 1494
1479 GluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerVal 498
1495 ACCTTGTTTAATGTTGCGAGCGCGCGCGCGTTCGCCGGTCCGCTCGATTGAGCAAG 1554
499 ThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLys 518
1555 GTCACACTGACGGTGGCCCTCTCCACCATCCAGAGTACTCGAAGACCTTCTTTGTC 1614
519 ValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheVal 538
1615 CTGCGCTCGCGGTAAAGCTCTTCTGCGGAGCGGACACAACTAAAGCGGGTACCTT 1674
539 LeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrPro 558
1675 TATAATTATAACACCACTGCTAGCGACCACTGCTTGTGAGAAATGCGCGCGGACCGG 1734
559 TyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGlyHisArg 578
1735 GTGCGTATTTCGACTTACACCACTAGCTGGGTGGTGGTCCCGCTCCATTCTCGGGTT 1794
579 ValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSerAlaVal 598
1795 GCGGTTTATAGCCCCCCTCTGCTAGCTAGCATTCGTTGAGGATACCTTGGACTACCTGCC 1854
599 AlaValLeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAla 618
1855 CGCGCCCACTACTTTGATGATTTCTGCCAGAGTGGCGCCCTTGGCTTCAGGCTGC 1914
619 ArgAlaHisThrPheAspAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCys 638
1915 GCTTTCAGTCTACTGCTGAGTTCAGCGCTTAAGATGAGTGGGTAAACTCGG 1974
639 AlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArg 658
1975 GAGTTG 1980
659 GluLeu 660

RESULT 6
US-09-769-066-25
Sequence 25, Application US/09769066
Patent No. US20020107360A1
GENERAL INFORMATION:
APPLICANT: Fuerst, Thomas R.
McAtee, C. Patrick
Yarbough, Patrick O.
Zhang, Yifan
TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/769,066
FILING DATE: 24-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/542,634
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 amino acids
TYPE: amino acid
STRANDEDNESS: Hepatitis E virus (Burma strain)
r52kDa, 58.1 kDa
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-769-066-25

Alignment Scores:
Pred. No.: 0 Length: 540
Score: 540.00 Matches: 540
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 81.82% Indels: 0
DB: 9 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-769-066-25 (1-540)

QY 334 GGGTGGTCCGGCCCATGACACCCGCGAGTGCCTGATGTGACATCCGCGCGCGCATC 393
Db 1 AlaValAlaProAlaHisAspThrProProValProAspValAspSerArgGlyAlaIle 20
QY 394 TTGCGCGCGCAGTATAACCTATCAACATCTCCCTTACCTCTTCCTGGCCACCGCACT 453
Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThr 40
QY 454 AACCTGGTCTTTATGCGCGCCCTCTTAGTCCGCTTTTACCCTTCAGGACGCGCAAT 513
Db 41 AsnLeuValLeuTyrAlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACCCATATATGCGCAGGAAGCTTCTAATTATTATGCGCGGTACGCCATCTCCATCTCATTC 573
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThr 80
QY 574 ATCCGTTTACCGCGCTGTCCTCCCAATGCTGTCGCGGTTCAGCCATCTCCATCTCATTC 633
Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCCACAGACACCAACCCGCGAGTCCGTTGATATGAATTAACCTCGACGGAT 693
Db 101 TrpProGlnThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120
QY 694 GTTCGTATTATTAGTCCAGCGCGCATAGCTCTGAGCTTGTGATCCCAAGTGAAGCCTTA 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTATCGTAACCAAGGCTGGCGCTCCGTCGAGACCTTCGGGTGGTGGTGGAGGAGGCT 813
Db 141 HisTyrArgAsnGlnGlyTrpArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCTGGTCTTGTATTGCTTGCATACATAGGCTCACTCGTAAATTCCTATATAATACA 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180
QY 874 CCTATACCGGTGCCCTCGGGCTGTGGACTTTGGCTTGGAGCTTGGAGTTTGCACACCTT 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200
QY 934 ACCCGCGGTAAACCAACCAATACGCGGGTCTCCCGTTATTTCAGCACTGCTGCCACCGCTT 993
Db 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
QY 994 CGTCGCGGTGCGGACGCGGACTGCGGAGCTCACACACGCGCTGCTATGAGTATGAG 1053


```
QY 694 GTTCGTAATTTAGTCCAGCCCGCATAGCTCTGAGCTTGATCCCAAGTCAGCGCTA 753
Db 121 ValArgIleuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTATCGTAACCAAGCGCTGGCGCTCGTCGAGACCTCTGGGTGGCTGAGGAGGAGCT 813
Db 141 HisTyrArgAsnGlnGlyTrpArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACTCTGGTCTGTATGCTTTCATACATAGGCTCACTCGTAAATTCCTATATAACA 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerLeuValAsnSerTyrThrAsnThr 180
QY 874 CCTATACCGTCCCTCGCGCTGTGGACTTTCCTTGAGCTTGAGCTTGCGCAACCTT 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsnLeu 200
QY 934 ACCCGCGTAACACCAATACGCGGTCTCCGTTATTCAGCACTCTCGCCACCGCTT 993
Db 201 ThrProGlyAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
QY 994 CGTCGGGTGGAGCGGACTCCGAGCTCACCAACCGCTCTACCGCTTTATGAAG 1053
Db 221 ArgArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLys 240
QY 1054 GACCTCTATTTACTAGTACTAATGTGTGGTGAGATCGCCGCGGATAGCCCTCAC 1113
Db 241 AspLeuTyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThr 260
QY 1114 CTGTTCAACTGCTGACACTCTGCTTGGCGCTCGCCAGACAGAAATTGATTCGTCGGCT 1173
Db 261 LeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAla 280
QY 1174 GGTGGCCAGCTGTTCTACTCCGTCCTCGCTTCCTCAGCAATGCGAGCGAGCTTAAG 1233
Db 281 GlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLys 300
QY 1234 TTGTATACATCTGTAGAGAAATGCTCAGCAGGATAGGGTATTGCAATCCGGATGACATT 1293
Db 301 LeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIle 320
QY 1294 GACCTCGGAGAAATCTGCTGTTATTCAGGATTAATGATAACCAACATGAACAGATCG 1353
Db 321 AspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArg 340
QY 1354 CGAGCGCTTCTCCAGCCCATCGCGCTTCTCTCTCTCTCGACTAATGATGCTT 1413
Db 341 ProThrProSerProAlaProSerArgProPheSerValIleArgAlaAsnAspValLeu 360
QY 1414 TGGCTCTCTCTCACCGCTCCGAGTATGACCACTTATGGCTCTTCGACTGGCCCA 1473
Db 361 TrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPro 380
QY 1474 GTTTATGTTCTGACTCTGTGACCTTGTTGTAATGTTGACCGCGCGGCGGCTTGC 1533
Db 381 ValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAla 400
QY 1534 CGGTCTCGATTGGACCAAGTCACTTGAAGTCCGCTCCCTCTCCACCATCCAGCAG 1593
Db 401 ArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGln 420
QY 1594 TACTCGAAGACCTCTTTTCTCTCGCTCCGCGTAACTCTCTTTCTGGAGGCGAGGC 1653
Db 421 TyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGly 440
QY 1654 ACAACTAAAGCGGGTACCTTTATTAATTAATACACCACTGCTAGGACCACTGCTTGC 1713
Db 441 ThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuVal 460
QY 1714 GAGATGCGCGCGGACCGGGTCGCTATTTCACCTTACACCACTAGCTGCGTCTGCT 1773
Db 461 GluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGly 480

1774 CCGGTCTCATTTCTCGGTTCGGTTTTAGCCCCCACTCTCGCTAGCATTCGTTGAG 1833
Db 481 ProValSerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuGlu 500
QY 1834 GATACCTTGAGCTACCTCGCGCCGCGCCATACCTTTTGATGATTTCTGCCAGAGTCCCGC 1893
Db 501 AspThrLeuAspTyrProAlaArgAlaHisThrPheAspAspPheCysProGluCysArg 520
QY 1894 CCGCTTGGCCTTCAG 1908
Db 521 ProLeuGlyLeuGln 525

RESULT 8
US-10-165-868-17
; Sequence 17, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Bradley, Daniel W.
; Twu, Jr-Shin
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; Filing Date: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; Filing Date: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; Filing Date: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; Filing Date: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; Filing Date: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; Filing Date: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
; Filing Date: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; Filing Date: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 436 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 9
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
```


US-10-165-868-17

Alignment Scores:

Pred. No.: 0 Length: 436
Score: 436.00 Matches: 436
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.06% Indels: 0
DB: 14 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-165-868-17 (1-436)

QY 673 AATTCAATAAAGCTCGACGATGTTTCGATTTTATAGTCCAGCCGCGATAGCCCTCAGCTT 732
DB 1 AsnSerIleThrSerThraspValArgIleLeuValGlnProGlyIleAlaSerGluLeu 20
QY 733 GTGATCCCAAGTAGCGCTACACTATCGTACCAAGGCTGGCGCTCCGTGGAGACCTCT 792
DB 21 VallileProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThrSer 40
QY 793 GGGTGGCTGAGGAGGAGGCTACCTCTGGCTTGTATGCTTTGTCATACATGCTCACTC 852
DB 41 GlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerLeu 60
QY 853 GTAAATTCCTATACATAACACCTATACCGCTGCCCTCGGCTGTGGACTTTTCCCTT 912
DB 61 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuAspPheAlaLeu 80
QY 913 GAGCTTGAGTTTCGCAACCTTTACCCCGGTACACCAATACCGGGTCTCCGGTTATTCC 972
DB 81 GluLeuGluPheArgAsnLeuThrProGlyAsnThrArgValSerArgTyrSer 100
QY 973 AGCACTGCTCGCCACCGCTTCGTCGGTGGGAGGAGGCTGGCGAGCTACACACGAG 1032
DB 101 SerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGluLeuThrThr 120
QY 1033 GCTGTACCGCTTATGAAGGACCTATTTTACTAGTACTAATGGTGGTGGTGGATC 1092
DB 121 AlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGlyValGlyGluLe 140
QY 1093 GGCCCGGGATGAGCCCTCACCCCTGTTCACCTTGCTGACACTCTCTGGCGGCTCGCG 1152
DB 141 GlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeuGlyGlyLeuPro 160
QY 1153 ACAGAAATGATTTGTCGGCTGGTGGCGAGCTGTCTACTCCGCTCCGCTTGTCTCAGC 1212
DB 161 ThrGluLeuIleSerSerAlaGlyGlnLeuPheTyrSerArgProValValSerAla 180
QY 1213 AATGCGAGCGGACGTTTAAAGTTGATACATCTGTAGAGATGCTCAGCAGATAGGGT 1272
DB 181 AsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGlnAspLysGly 200
QY 1273 ATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGTGGTTATTTCAGGATTATGAT 1332
DB 201 IleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIleGlnAspTyrAsp 220
QY 1333 AACCAATGATACAGATCGCGGAGCTTCTCAGCCCACTCGGCGCTTCTCTGTC 1392
DB 221 AsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerVal 240
QY 1393 CTTTCAGCTTAATGATGTCTTGGCTCTCTCAGCGCTCGGAGTATGACCACTCCACT 1452
DB 241 LeuArgAlaAsnAspValLeuIlePheSerLeuThrAlaAlaGluTyrAspGlnSerThr 260
QY 1453 TATGGCTCTTCGACTGGCCAGTTTATGTTTCTGACTCTGTGACCTTGGTTTAAATGTGGC 1512
DB 261 TyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeuValAsnValAla 280
QY 1513 ACCGGCGGAGCGGCTTCCCGGCTGGCTCGATGAGCAAGGTCACACTTACCGGTCGC 1572
DB 281 ThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArg 300
QY 1573 CCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTCCGCTCCGCGTAG 1632

DB 301 ProLeuSerThrIleGlnGlnTyrSerLysThrPheValLeuProLeuArgGlyLys 320
QY 1633 CTCTCTTCTGGAGGAGGAGCACAATAAGCCGGGTACCTTATAATATACACCACT 1692
DB 321 LeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThr 340
QY 1693 GCTAGCGACCAACTGCTTGTGAGAATCGCGCGGAGCACCAGGTGCTATTTCACCTTAC 1752
DB 341 AlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyr 360
QY 1753 ACCACTAGCCTGGGTGGTGGTCCCGCTCTCCATTTCTGCGGTGGCGTTTATAGCCCCCAC 1812
DB 361 ThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaValLeuAlaProHis 380
QY 1813 TCTGGCTAGCATGCTTGGAGTACCTTGGACTACCTCGGCTCCGCGCGCCATACCTTTGAT 1872
DB 381 SerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAlaHisThrPheAsp 400
QY 1873 GATTCTCCCGCAGAGTCCGCCCTTGGCTTCCAGGGCTCGCTTTCAGTCTACTCTGTC 1932
DB 401 AspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPheGlnSerThrVal 420
QY 1933 GCTGAGCTTCAGCGCTTAAAGTGAAGTGGGTAAACTCGGGAGTTG 1980
DB 421 AlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 436

RESULT 9

US-10-381-770-1
; Sequence 1, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS
; FILE REFERENCE: IEC010037PCT
; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 660
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-1

Alignment Scores:
Pred. No.: 0 Length: 660
Score: 408.00 Matches: 508
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 61.82% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-1 (1-660)

QY 289 GCCTCAGCTCGTAGACTACACAGCTGGGCGCGCGCTAACCGCGTCTCGGCGC 348
DB 97 AlaSerArgArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAla 116
QY 349 CATGACACCCCGCAGTGCCTGATGCTGACTCCCGCGCGCCATCTTGGCGCGGAGTAT 408
DB 117 HisAspThrProProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyr 136
QY 409 AACCTATCAACATCTCCCTTACCTTTCGTTGGCCACCGGCACTAACTGGTCTTTAT 468
DB 137 AsnLeuSerThrSerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyr 156
QY 469 GCCGCGCCCTCTAGTCCGCTTTTACCCCTTACGAGGAGGAGCCCAATACCATATATGCC 528
DB 157 AlaAlaProLeuSerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAla 176
QY 529 ACGGAAGCTTCTTAATTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588

Db 177 ThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgPro 196
QY 589 CTGGTCCCAATGCTGTGGCGGTAGCCATCTCCATCTCATCTGCGGCACAGACACC 648
Db 197 LeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThr 216
QY 649 ACCACCCGAGTCCGTTGATGAATTCATTAACCTCGACGGATGTCGTATTTAGTC 708
Db 217 ThrThrProThrSerValAspMetAsnSerIleThrSerThrAspValArgIleLeuVal 236
QY 709 CAGCCCGCATAGCTCTGAGCTTGTGATCCCAAGTACAGCCCTACACTATCGTAACCA 768
Db 237 GlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGln 256
QY 769 GGTGGCGCTCCCTCGAGACTCTGCGGTGGCTGAGAGAGAGCTACCTCTGCTGTGTT 828
Db 257 GlyTyrArgSerValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuVal 276
QY 829 ATGCTTTGCATACATGCTCACTCGTAAATTCCTACTATAATACACCTATACCGGTGCC 888
Db 277 MetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAla 296
QY 889 CTCGGCTGTGAGACTTGCCTTGAGCTTGTGAGTTTCGCAACCTTACCCCGGTAAACACC 948
Db 297 LeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThr 316
QY 949 AATACGGGTCTCCCGTTATTCAGACTGTGCGGACCGCTTGTGCGGTGGGAC 1008
Db 317 AsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeuArgGlyAlaAsp 336
QY 1009 GGGACTGCCGAGCTCACACACCGCTGCTACCGCTTATTAAGAGACCTCTATTTACT 1068
Db 337 GlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetIysAspLeuTyrPheThr 356
QY 1069 AGTACTAATGGTGTGCGTGAGATCGCGCGGATAGCCCTCAACCTGTTCACCTTGCT 1128
Db 357 SerThrAsnGlyValGlyGluIleGlyArgGlyAlaLeuThrLeuPheAsnLeuAla 376
QY 1129 GAACTGTGTTGGCGCTGCGGACAGATTAATGTTGCTGCGCTGCGGACCTGTTTC 1188
Db 377 AspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPhe 396
QY 1189 TACTCCGCTCCGCTGCTCAGCAATGGCGAGCGACTGTTAAGTTGTATACATCTGTA 1248
Db 397 TyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyrThrSerVal 416
QY 1249 GAGATGCTCAGCAGATTAAGGTATTGCAATCCCGCATGACATTCACCTCGGAGATCT 1308
Db 417 GluAsnAlaGlnAspIysGlyIleAlaIleProHisAspIleAspLeuGlyGluSer 436
QY 1309 CGTGTGTTATTGAGATTAATGATACCAACATGAACAGATCGCGGACGCTTCTCCA 1368
Db 437 ArgValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPro 456
QY 1369 GCCCATCGCGCTTCTCTGCTTCGAGCTAATGATGCTGTTGGCTCTCTCTCACC 1428
Db 457 AlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThr 476
QY 1429 GTGCGGATGATGACAGTCCACTTATGGCTTTCGACTGCGGACGCTTATGTTCTGAC 1488
Db 477 AlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAsp 496
QY 1489 TCTGTGCTGCTGTTAATGTTGCGCGCGCGGCGCTGCGGCTGCGCTGCTGCTGCTG 1548
Db 497 SerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrp 516
QY 1549 ACCAGGTCACTTACAGCTGCGGCTTCCACCATCCAGCAGTACTCGAAGACCTTC 1608
Db 517 ThrIysValThrLeuAspGlyArgProLeuSerThrIleGlnIleTyrSerIysThrPhe 536
QY 1609 TTGTCTGCTGCGCTGCGGCTGAGCTCTCTTCTGGAGGCGGACCACTAAGCGGG 1668

Db 537 PheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGly 556
QY 1669 TACCCTTATAATATTAACACCACTGCTAGGACCAACTGCTTCTGAGAAATGCGCGGG 1728
Db 557 TyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGly 576
QY 1729 CACCGGTGCTGCTATTTCCACTTACACACTAGCTGCTGCTGCTGCTGCTGCTGCT 1788
Db 577 HisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSer 596
QY 1789 GCGTTGCCGCTTTAGCCCCCACTCT 1815
Db 597 AlaValAlaValLeuAlaProHisSer 605
RESULT 10
US-09-769-066-17
; Sequence 17, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; McAttee, C. Patrick
; Yarbrough, Patrice O.
; Zhang, Yifan
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09769,066
; FILING DATE: 24-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/542,634
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 amino acids
; TYPE: amino acid
; STRANDEDNESS: Hepatitis E Virus (Burma strain) SG3
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-769-066-17
Alignment Scores:
Pred. No.: 2,51e-293 Length: 327
Score: 327.00 Matches: 327
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
DB: 9 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-769-066-17 (1-327)

QY 1000 GGTGGGACGGAGCTGCGAGCTCACACCACGGCTGTACCGCTTTTATGAGGACCTC 1059
Db 1 GlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetIysAspLeu 20
QY 1060 TATTTTACTAGTACTAATGGTTCGGTGAGATCGCGCGGGATACCCCTCACCTGTTC 1119
Db 21 TyrPheThrSerThrAsnGlyValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPhe 40
QY 1120 AACCTTGCTGACACTCTGCTGGCGCCCTGCGACAGAAATTATTCGTGCGCTGTGGC 1179
Db 41 AsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGly 60
QY 1180 CAGCTGTTCTACTCCCGTCCCGTGTCTCAGCCCAATGGCGAGCGCACTGTTAAGTTGTAT 1239
Db 61 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 80
QY 1240 ACATCTGTAGAGATGCTCAGCAGGATAAGGTATTGCAATCCCGCATGACATTGACCTC 1299
Db 81 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 100
QY 1300 GGAGATCTCGTGTGGTATTATTCAGGATATGATACCAACATGACCAAGATCGCGGAGC 1359
Db 101 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 120
QY 1360 CCTTCTCCAGCCCCATCGCGCCCTTCTCTGCTCTCGAGCTAATGATGCTGTTGGCTC 1419
Db 121 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 140
QY 1420 TCTCTCACCGTCCGAGTATCACCAGTCCACTTATGGCTCTTCGACTGGCCAGTTTAT 1479
Db 141 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 160
QY 1480 GTTTCGTGACTCTGTGACCTTGTTAATGTTGGACCGCGCGCAGCGCTTCCCGGTGC 1539
Db 161 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 180
QY 1540 CTCGATTGACCAAGCTCACACTTCAGCGTCCCGCTCTCCACCATCCAGCAGTACTCG 1599
Db 181 LeuAspThrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 200
QY 1600 AAGACCTTTTGTCTGCTGCGCTCCGCGTAAAGCTCTCTTCTGGGAGGACGACAACT 1659
Db 201 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 220
QY 1660 AAAGCGGGTACCTTATTAATTAACACCATGCTAGGACCACTGCTTTCAGAAAT 1719
Db 221 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 240
QY 1720 GCCGCGCGGACCGGTGCTGCTATTTCCACTTACACCATAGCTGGGTGCTGCTCCCGTC 1779
Db 241 AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 260
QY 1780 TCCATTCTCGGGTTCGGTTTTAGCCCCCTAGCTGCGCTAGCATGCTCAGGATACC 1839
Db 261 SerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuGluAspThr 280
QY 1840 TTGGACTACCTGCGCGGCCCATACTTTTGATGATTCTGCGCAGAGTCCGCCCTT 1899
Db 281 LeuAspTyrProAlaArgAlaHisThrPheAspAspPheCysProGluCysArgProLeu 300
QY 1900 GGCCTTCAGGGTGGCTTTCAGTCTACTGTCGCTAGCTTCAGCGCTTAAAGATGAAG 1959
Db 301 GlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuIysMetIys 320
QY 1960 GTGGGTAAACTCGGAGTTG 1980
Db 321 ValGlyLysThrArgGluLeu 327

RESULT 11

US-10-165-868-15

; Sequence 15, Application US/10165868

; Publication No. US20030143241A1

GENERAL INFORMATION:
APPLICANT: Reyes, Gregory R.
Bradley, Daniel W.
Tsu, Jr-Shin
Purdy, Michael A.
Tam, Albert W.
Krawczynski, Krzysztof Z.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/165,868
FILING DATE: 06-Jun-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/870,985A
FILING DATE: 20-APRIL-1992
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0093.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 327 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: SG3, BURMA, FIGURE 9
SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-165-868-15
Alignment Scores:
Pred. No.: 2,51e-293 Length: 327
Score: 327.00 Matches: 327
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
DB: 14 Gaps: 0
US-09-851-410a-6_COPY_5147_7129 (1-1983) x US-10-165-868-15 (1-327)

QY 1000 GGTGGGACGGAGCTGCGAGCTCACACCACGGCTGTACCGCTTTTATGAGGACCTC 1059

Db 1 GlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetIysAspLeu 20

QY	1060	TATTTTACTAGTACTAATAGTGTCGTTGAGATCGCGCGGGATGAGCCCTCACCTCTGTC	1119
Db	21	TyrPheThrSerThrAsnGlyValGlyIleGlyArgGlyIleAlaLeuThrLeuPhe	40
QY	1120	AACCTTCTGACACTCTGCTTGGCGGCTCGCAGACAGAAATGATTTCTGTCGGCTGGTGGC	1179
Db	41	AsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGly	60
QY	1180	CAGCTGTCTTACTCTCCGTCCTCGTGTCTTCACCAATGCGAGCCGACTGTTAACTTGTAT	1239
Db	61	GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr	80
QY	1240	ACATCTGPAGAAATGCTCAGCAGGATAAGGTTATGCAATCCGCATGACATGACCTC	1299
Db	81	ThrSerValGluAsnAlaGlnAspIysGlyIleAlaIleProHisAspIleAspLeu	100
QY	1300	GGAAATCTCTGTGCTTATTCAGGATTATCAACCAATGACCAAGATCGGCCGACG	1359
Db	101	GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr	120
QY	1360	CTTCTCCAGCCCATCGCGCCTTCTCTCTCTCTCGAGCTAATGATGCTTTGGCTC	1419
Db	121	ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu	140
QY	1420	TCTCTCACCGTGGCAGTATGACACAGTTCACCTTATGGCTCTTCGACTGGCCCACTTAT	1479
Db	141	SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr	160
QY	1480	GTTTCTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGAGCGGTGCCCGTCG	1539
Db	161	ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer	180
QY	1540	CTCGATTCCGACCAAGGTCACACTTCAGCGTCCGCCCTCTCCACATCCACGACGTACTCG	1599
Db	181	LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnTyrSer	200
QY	1600	AAGACTTCTTTGTCTCCGCTCCGCGTAAAGTCTCTTCTTCGGAGCGCAGGCACAACT	1659
Db	201	LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr	220
QY	1660	AAAGCCGGTACCTTATAATTATACACCACTGCTAGCGACCACTGCTGTGTCGAGAT	1719
Db	221	LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsn	240
QY	1720	GCCGCGCGGACCGGTCGCTATTTCCACTTTACACCACTAGCTCGGTGGTCTGTCCTGTC	1779
Db	241	AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal	260
QY	1780	TCCATTCTGCGGTTGCGCTTTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGATACC	1839
Db	261	SerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuGluAspThr	280
QY	1840	TTGGACTACCTGCGCGGCCCATACTTTTATGATTTCTGCCCAGAGTCCGCGCCCTT	1899
Db	281	LeuAspTyrProAlaArgAlaHisThrPheAspAspPheCysProGluCysArgProLeu	300
QY	1900	GCCCTTCAGGGCTCGCCTTTCAGTCTACTCTGCTGACTTCAGCTTCAGCGCTTAAAGATGAAG	1959
Db	301	GlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuIysMetIys	320
QY	1960	GTGGTAAACCTCGGAGTTG	1980
Db	321	ValGlyLysThrArgGluLeu	327

RESULT 12.

RESULT 12
US-10-381-770-2

US-10-381-770-2
: Sequence 2: Application US/10381770

; sequence 2, Application US/I038
 : Publication No. US20040052813A1

; Publication No. US20
; GENERAL INFORMATION:
; GENERAL INFORMATION:

; GENERAL INFORMATION:
: APPLICANT: YANG SHENG TANG company. Ltd

APPLICANT: YANG SHENG TANG company, Ltd.
TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS. THE VACCINE COMPOSITION

; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE
: TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS

; TITLE OF INVENTION: COMPOSITE
: FILE REFERENCE: IFC010037PCT

```

, CURRENT APPLICATION NUMBER: US/10/381,770
,
, CURRENT FILING DATE: 2003-07-16
,
, NUMBER OF SEQ ID NOS: 12
,
, SOFTWARE: PatentIn version 3.1
,
, SEQ ID NO 1
,
, LENGTH: 214
,
, TYPE: PRT
,
, ORGANISM: hepatitis E virus
,
, US-10-381-770-2

```

Alignment Scores:

Alignment Scores:		
Pred. No.:	5.37e-185	214
Score:	210.00%	210
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0
Query Match:	31.82%	0
DB:	12	0
	Gaps:	0
	Indels:	0
	Mismatches:	0
	Conservative:	0
	Matches:	210
	Length:	214

US-09-851-410A-6 COPY 5147 7129 (1-1983) X US-10-381-770-2 (1-214)

Qy	1180	CAGCTGTTCTACTCCCTCCCGTTGCTCTAGCCAAATCGCGAGCGCACTGTTAAGTTGAT	1239
Db	2	GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr	21
Qy	1240	ACATCTCTAGAGAAATGCTCAGCAGGATAAGGTATTGCAATCCCGCATGACATTGACCTC	1299
Db	22	ThrSerValGluAsnAlaGlnGlnAspLysGlyLeuAlaIleProHisAspIleAspLeu	41
Qy	1300	GGAAATCTCGTGTGGTATTTCAGGATTATGATAACCAACATGAACAGATCGCCGACG	1359
Db	42	GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr	61
Qy	1360	CTTCTCCAGCCCAATCGCCCTTTCTCTGTCTCTCAGCTAATGATGTCTTTGGCTC	1419
Db	62	ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu	81
Qy	1420	TCTCTCACCGTGCAGAGTAGCCAGCTGCACATTATGGCTCTTCGACTGGCCCAAGTTTAT	1479
Db	82	SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr	101
Qy	1480	GTTTCTGACTCTGTGACTGTTGGTTAAATGTTGCGACCGCGCGAGGCGGTTGCCGGTGG	1539
Db	102	ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer	121
Qy	1540	CTCATTTGGACAGGTGCACATGTAGCGTGCGCCCTCTCCACCATCCAGCACTACTCG	1599
Db	122	LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer	141
Qy	1600	AAGACCTTCTTTGTCTGCCGCTCCGGGGTAAGCTCTCTTTCTGGGAGCGACGCAACT	1659
Db	142	LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr	161
Qy	1660	AAAGCGGGTACCTTATAATTATACACCACTGCTAGGACCAACTGCTTGTGAGAAAT	1719
Db	162	LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsn	181
Qy	1720	GCCGCGCGGACCGGGTCGCTATTTCACATTACACCACTAGCCTGGGTGCTGGTCCGCTC	1779
Db	182	AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal	201
Qy	1780	TCCATTCTTCGGTTCGGTTTTAGCCCCC	1809
Db	202	SerIleSerAlaValAlaValLeuAlaPro	211

RESULT 13

US-10-381-770-12

US-10-381-770-12
; Sequence 12, Application US/10381770

Sequence 12, Application US/103
Publication No. US20040052813A1

; PUBLICATION NO. 0320
: GENERAL INFORMATION:
: GENERAL INFORMATION:

; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.

APPLICANT: YANG SHENG YANG COMPANY, LTD.
TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPO

;
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, AND
; TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS

; TITLE OF INVENTION: COMPRIS
: FILE REFERENCE: IEC010037PCT

; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 232
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-12

Alignment Scores:
Pred. No.: 5,32e-185 Length: 232
Score: 210.00 Matches: 210
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.82% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-12 (1-232)

QY 1180 CAGCTGTTCTACTCCCGTCCCGTGTCTCAGCAATGGCGAGCGGACTGTAAAGTTGTAT 1239
Db 23 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 42
QY 1240 ACATCTGTAGAGATGCTCAGCAGGATAAGGTATTGCAATCCCGCATGACATTCACCTC 1299
Db 43 ThrSerValGluAsnAlaGlnAspLysGlyLeuAlaIleProHisAspLeu 62
QY 1300 GGAGATCTCGTGTGTTATTCAGGATTATGATAACCAATGACCAATGACGCGGAGC 1359
Db 63 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 82
QY 1360 CCTTCTCCAGCCCCATCGCGCCCTTCTCTGCTCCGAGTCAATGATGCTTTGGCTC 1419
Db 83 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrPleu 102
QY 1420 TCTCTCAGCGTCCGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCAGTTAT 1479
Db 103 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 122
QY 1480 GTTCTGACTCTGTGACCTTGTATTAATGTTGGAGCGGCGCGGCGGCTGCCCGGTG 1539
Db 123 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 142
QY 1540 CTCGATTGACCAAGTCACTTCAGCGTCCCGCTCTCCACCATCCAGCAGTACTCG 1599
Db 143 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 162
QY 1600 AAGACCTTCTTGTCTCGCGCTCCCGGTAAGTCTCTTTCTGGAGGACGACAACT 1659
Db 163 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 182
QY 1660 AAAGCGGGTACCTTATATTAATTAACCACTGCTAGGACCACTGCTTCTTCAGAA 1719
Db 183 LysAlaGlyTyrProTyrAsnTyrAsnThrAlaSerAspGlnLeuValGluAsn 202
QY 1720 GCCCGCGGCGACCGGTCGCTATTTCCACTTACACCACTAGCTGGTCCCGTC 1779
Db 203 AlaAlaGlyHisArgValAlaIleSerThrThrThrSerLeuGlyAlaGlyProVal 222
QY 1780 TCCATTCTCGGGTTCGGTTTTAGCCCC 1809
Db 223 SerIleSerAlaValAlaValLeuAlaPro 232

RESULT 14

US-10-381-770-4
; Sequence 4, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS
; TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
; FILE REFERENCE: IEC010037PCT

; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 389
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-4

Alignment Scores:
Pred. No.: 9,01e-176 Length: 389
Score: 200.00 Matches: 200
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.30% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-4 (1-389)

QY 1240 ACATCTGTAGAGATGCTCAGCAGGATAAGGTATTGCAATCCCGCATGACATTCACCTC 1299
Db 143 ThrSerValGluAsnAlaGlnAspLysGlyLeuAlaIleProHisAspLeu 162
QY 1300 GGAGATCTCGTGTGTTATTCAGGATTATGATAACCAATGACCAATGACGCGGAGC 1359
Db 163 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 182
QY 1360 CCTTCTCCAGCCCCATCGCGCCCTTCTCTGCTCCGAGTCAATGATGCTTTGGCTC 1419
Db 183 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrPleu 202
QY 1420 TCTCTCAGCGTCCGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCAGTTAT 1479
Db 203 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 222
QY 1480 GTTCTGACTCTGTGACCTTGTATTAATGTTGGAGCGGCGCGGCGGCTGCCCGGTG 1539
Db 223 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 242
QY 1540 CTCGATTGACCAAGTCACTTCAGCGTCCCGCTCTCCACCATCCAGCAGTACTCG 1599
Db 243 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 262
QY 1600 AAGACCTTCTTGTCTCGCGCTCCCGGTAAGTCTCTTTCTGGAGGACGACAACT 1659
Db 263 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 282
QY 1660 AAAGCGGGTACCTTATATTAATTAACCACTGCTAGGACCACTGCTTCTTCAGAA 1719
Db 283 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 302
QY 1720 GCCCGCGGCGACCGGTCGCTATTTCCACTTACACCACTAGCTGGTCCCGTC 1779
Db 303 AlaAlaGlyHisArgValAlaIleSerThrThrThrSerLeuGlyAlaGlyProVal 322
QY 1780 TCCATTCTCGGGTTCGGTTTTAGCCCCCGGCTAGCTGGTCCGTTGAGGATACC 1839
Db 323 SerIleSerAlaValAlaValLeuAlaProHisSerAlaLeuAlaLeuGluAspThr 342

RESULT 15
US-10-381-770-3
; Sequence 3, Application US/10381770
; Publication No. US20040052813A1
; GENERAL INFORMATION:
; APPLICANT: YANG SHENG TANG company, Ltd.
; TITLE OF INVENTION: THE POLYPEPTIDE FRAGMENT OF HEPATITIS E VIRUS, THE VACCINE COMPOS
; TITLE OF INVENTION: COMPRISING SAID FRAGMENTS AND THE DIAGNOSTIC KITS
; FILE REFERENCE: IEC010037PCT
; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1

```

; SEQ ID NO 3
; LENGTH: 194
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-3

Alignment Scores:
  Pred. No.:      194
  Score:          190.00
  Percent Similarity: 100.00%
  Best Local Similarity: 100.00%
  Query Match:    28.79%
  Indels:         0
  Gaps:           0
  DB:             12

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-3 (1-194)

QY 1240 ACATCTGTAGAGAAATCCTCAGCAGGATTAAGGGTATTGCAATCCCGCATGACATTGACCTC 1299
DB 2 ThrSerValGluAsnAlaGlnGlnAspLysGlyLeAlaIleProHisAspIleAspLeu 21

QY 1300 GGAGAAATCTCGTGTGTTATTCAGGATTATGATAACACATGACACAGATCGCCGACG 1359
DB 22 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 41

QY 1360 CCTTCTCCAGCCCATCGCCGCTTCTCTGCTTCGAGCTAATGATGCTTTGGCTC 1419
DB 42 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 61

QY 1420 TCTCTCACCGCTCCCGAGTATGACCATTCACCTTATGGCTCTTTCGACTGCCCCAGTTTAT 1479
DB 62 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 81

QY 1480 GTTTCGTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGCAGCGCTTGCCCGTGC 1539
DB 82 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 101

QY 1540 CTGATTTGGACCAAGGTCACACTTGAGCGTCCGCCCTCTCCACCATCCAGCAGTACTCG 1599
DB 102 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 121

QY 1600 AAGACCTTCTTCTGCTCCGCGCTCCGCGGTAAAGCTCTCTTCTGGGAGGCGAGGCACAAC 1659
DB 122 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 141

QY 1660 AAAGCCGGGTACCTTTATATATATATACACACTGCTAGCAGCACCACTGCTTGTGAGAA 1719
DB 142 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsn 161

QY 1720 GCCGCGCGGCGACCGGTCCGCTATTTCACCTTACACCACTAGCCTGGGTGCTGGTCCGCTC 1779
DB 162 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 181

QY 1780 TCCATTTCTCGGTTGCGGTTTATAGCCCCC 1809
DB 182 SerIleSerAlaValAlaValLeuAlapro 191

```

Search completed: August 16, 2004, 14:29:20
Job time : 122.324 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:00:26 ; Search time 97,3236 Seconds
(without alignments)
12792.733 Million cell updates/sec

Title: US-09-851-410A-6_COPY_5147_7129

Perfect score: 3649

Sequence: 1 ATGCGCCCTGGCGTATTTT.....GTAAGAACTCGGAGTGTAG 1983

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 2585610

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlip
-Q=/cgm2_1/USPTO_spool_p/US09851410/runat_16082004_125455_8610/app_query.fasta_1.4238
-DB=Published Applications AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum2
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09851410@cgm2_1 27 @runat_16082004_125455_8610
-NCFU=6 -ICFU=3 -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:**

1: /cgm2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgm2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgm2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgm2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgm2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgm2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgm2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgm2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgm2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgm2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgm2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgm2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgm2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgm2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgm2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgm2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgm2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgm2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	3408	93.4	660	9	US-09-769-066-13	Sequence 13, Appl
2	3408	93.4	660	10	US-09-851-410-8	Sequence 8, Appl
3	3408	93.4	660	14	US-10-165-868-19	Sequence 19, Appl
4	3391	92.9	660	12	US-10-381-770-1	Sequence 1, Appl
5	3383	92.7	660	15	US-10-257-044-1	Sequence 14, Appl
6	3221	88.3	660	9	US-09-769-066-14	Sequence 1, Appl
7	3194.5	87.5	659	14	US-10-165-868-20	Sequence 20, Appl
8	3179	87.1	674	16	US-10-239-090A-57	Sequence 57, Appl
9	3176	87.0	660	16	US-10-239-090A-51	Sequence 51, Appl
10	3174	87.0	660	16	US-10-239-090A-53	Sequence 53, Appl
11	3166	86.8	674	16	US-10-239-090A-55	Sequence 55, Appl
12	3160	86.6	660	10	US-09-468-147-92	Sequence 92, Appl
13	3160	86.6	660	12	US-10-319-745-92	Sequence 92, Appl
14	3151	86.4	660	10	US-09-468-147-167	Sequence 167, App
15	3151	86.4	660	12	US-10-319-745-167	Sequence 167, App
16	2813	77.1	549	9	US-09-769-066-15	Sequence 15, Appl
17	2769	75.9	540	9	US-09-769-066-25	Sequence 25, Appl
18	2694	73.8	525	9	US-09-769-066-27	Sequence 27, Appl
19	2684	73.6	549	9	US-09-769-066-16	Sequence 16, Appl
20	2644	72.5	540	9	US-09-769-066-26	Sequence 26, Appl
21	2569	70.4	525	9	US-09-769-066-28	Sequence 28, Appl
22	2232	61.2	436	14	US-10-165-868-17	Sequence 17, Appl
23	2098.5	57.5	435	14	US-10-165-868-18	Sequence 18, Appl
24	1934.5	53.0	459	10	US-09-468-147-206	Sequence 206, App
25	1934.5	53.0	459	10	US-09-468-147-207	Sequence 207, App
26	1934.5	53.0	459	12	US-10-319-745-206	Sequence 206, App
27	1934.5	53.0	459	12	US-10-319-745-207	Sequence 207, App
28	1683.5	46.1	389	12	US-10-381-770-4	Sequence 4, Appl
29	1677	46.0	327	9	US-09-769-066-17	Sequence 17, Appl
30	1677	46.0	327	14	US-10-165-868-15	Sequence 15, Appl
31	1601	43.9	338	10	US-09-468-147-199	Sequence 199, App
32	1601	43.9	338	10	US-09-468-147-200	Sequence 200, App
33	1601	43.9	338	12	US-10-319-745-199	Sequence 199, App
34	1601	43.9	338	12	US-10-319-745-200	Sequence 200, App
35	1594	43.7	327	9	US-09-769-066-18	Sequence 18, Appl
36	1594	43.7	327	14	US-10-165-868-16	Sequence 16, Appl
37	1592	43.6	327	10	US-09-468-147-176	Sequence 176, App
38	1592	43.6	327	12	US-10-319-745-176	Sequence 175, App
39	1581	43.3	327	10	US-09-468-147-175	Sequence 175, App
40	1581	43.3	327	12	US-10-319-745-175	Sequence 175, App
41	1352	37.1	276	10	US-09-468-147-189	Sequence 189, App
42	1352	37.1	276	12	US-10-319-745-189	Sequence 189, App
43	1334.5	36.6	606	14	US-10-029-840-6	Sequence 6, Appl
44	1084	29.7	232	12	US-10-381-770-12	Sequence 12, Appl
45	1076	29.5	214	12	US-10-381-770-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-769-066-13
; Sequence 13, Application US/09769066
; Patent No. US20020107360A1
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; McAtee, C. Patrick
; Yarbrough, Patrice O.
; Zhang, Yifan

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/769,066
FILING DATE: 24-Jan-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/542,634
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0293.30

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 660 amino acids

TYPE: amino acid

STRANDEDNESS: Hepatitis E Virus (Burma strain)

ORF-2

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-09-769-066-13

Alignment Scores:

Pred. No.: 5,91e-243 Length: 660
Score: 3408.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.40% Indels: 0
DB: 9 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-769-066-13 (1-660)

QY 1 ATGCGCCCTCGCCATTTTGTGCTCTCATGTTTTCCTATGTCGCCCGGCCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCCCGGTACGCGTCTGGCGCGTCTGGCGCGCGAGCGGGTTCGGCGGTGT 120
Db 21 ProProGlyGlnProSerGlyArgArgGlyArgSerGlyGlySerGlyGly 40
QY 121 TTCTCGGGTACCGGTTGATTCAGCCCTTCGCAATCCCTATATTCATCCCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProIleHisProThrAsn 60
QY 181 CCTTCGCGCCCGATGTCACCGCTGCGCGCGGGCTGGACCTCGTGTTCGCAACCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGTCGCTGGCGTGGACAGGCCCGCGCGCGCGCGCGTTCACGTGCT 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArg 100
QY 301 AGACCTACACAGTGGGCGCGCGCTAACCGCGTCTCGCGCCCATCACACCCCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CGAGTGCCTGATGTGACATCCCGCGCGCCATCTTCGCGCGCGCAGTAACTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnThrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCGCGCCCGACCTAACCTGGTTCCTTATGCGCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyraAlaProLeu 160
QY 481 AGTCGCGCTTTACCTTCAGGACCGCACCAATACCATATATATGCGCAGGAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGlualaSer 180
QY 541 AATTATGCCACGATACCGGGTTGCCCGTGCCACAAATCCGTTACCGCGCGTGTCCCAAT 600

Db 181 AsnTyrAlaGlnTyrArgValAlaAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTGCGGGTTAGCCATCTCCATCTCATCTCTGGCCACACACACACACCGCCGAG 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATACCTACCTGAGGATGTTCTGTTATTTAGTCCAGCCGCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGCTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCCTCACTGTAATTCCTATATACACCTATACACCTGCGGTCCTCGGGCTGTTG 900
Db 281 HisGlySerLeuValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGTTATTCAGCACTGCTCGCCACCGCTTCGTCGGGTGCGGAGCGGAGTCCGCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGCTGCTACCGCTTTATGAAGGACCTCTATTTACTAGTACTAATGCT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGTGAGATCGCGCGGGATAGCCCTACCCCTGTTCAAACCTTCTGCTGACACTCTGCTT 1140
Db 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCGACAGATGATTTCTGCGGTGCTGGCGCGCGCTGCTCTACTCCGCTCCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATCGCGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGTPATTGCAATCCCGCATGACATGACATGACCTCGGAGAAATCTCGTGTGTTAT 1320
Db 421 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 440
QY 1321 CAGGATTTATGATAACCAACATCAAGATCGCGGAGCGCTTCTCCAGCCCATCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CTTTCTCTGCTCTGCTGAGCTAAATGATGCTTTTGGCTCTCTCTCACCGCTCCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCATTATGGCTCTTCGACTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTATGTTGGACCGCGCGCGCGCGCTGCTGCGGTGCTCGATTGACCAAGTTCACA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTGCGGTGCGCCCTCTCCACCATCCAGCAGTACTGAAGACCTCTTTTCTCTGCGCG 1620
Db 521 LeuAspGlyArgProLeuSerThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAACTCTCTTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680

Db 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGCTAGCGACCACTGCTCTCCAGATGCGCGGGACCGGGTCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCACACTTACACACTAGCTGCGTGGTCTGCTCCGCTCTCCATTTCTTCGCGTTGCCGTT 1800
Db 581 IleSerThrTyrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTGCGCTAGCACTGCTTGCAGATACCTTGGACTACCTGCCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaLeuLeuGluAspThrLeuAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTCCGCGCCCTTGGCCTTCAGGGCTGGCCTTC 1920
Db 621 HisThrPheAspPheCysProGluCysArgProLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTGAGCTTCAGCGCCTTAAGATGAGGTGGGTAAAACTCGGGAGTTG 1980
Db 641 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGluLeu 660
RESULT 2
US-09-851-410-8
; Sequence 8, Application US/09851410
; Publication No. US20030124510A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R
; Yarbough, Patricia O
; Bradley, Daniel W
; Krawczynski, Krzysztof Z
; Tam, Albert
; FTY, Kirk E
; TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-May-2001
; PRIORITY DATE: 07-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/128,275
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/681,078
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,888
; FILING DATE: 05-APR-1990
; APPLICATION NUMBER: US 07/420,921
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: US 07/367,486
; FILING DATE: 16-JUN-1989
; APPLICATION NUMBER: US 07/336,672
; FILING DATE: 11-APR-1989
; APPLICATION NUMBER: US 07/208,997
; FILING DATE: 17-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 4600-0183.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 324-0880
; TELEFAX: (650) 324-0960

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-851-410-8
Alignment Scores:
Pred. No.: 5,91e-243 Length: 660
Score: 3408.00 Matches: 660
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.40% Indels: 0
DB: 10 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-851-410-8 (1-660)
QY 1 ATGCGCCCTCGGCCTATTTTGTCTGCTCTCATGTTTTCCTATGCTGCTGCGCGCA 60
Db 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGGTACAGCGCTCTGGCGCGCTGCTGCGCGCGCGCACGCGCGGTTCGCGCGGTG 120
Db 21 ProProGlyGlnProSerGlyArgArgArgGlyArgArgSerGlyGlySerGlyGlyGly 40
QY 121 TTCTGGGGTGACCGGGTTGATTCAGCCCTTCGCAATCCCTATATTCATCAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCTTCGCGCCCGGATGTCACCGCTGCGCGCGCGGTGACCTGCTGTTGCGCAACCCGCC 240
Db 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCCGCTTGGCGGTGACCGCGCGCGCGCGCGCGGTTCGCTCAGCTG 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaValAlaSerArgArg 100
QY 301 AGACTACACAGCTGGGGCGCGCGCTAACCGCGGTGCTGCGCGCGCATGACACCCCG 360
Db 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTCTGATGTCGACTCCCGCGCGCGCGCATCTTGGCGCGCGAGTATAACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCCGTGGCCACCGCGCACTAACCTGGTTCCTTTATGCGCGCGCTCT 480
Db 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGGCTTTTACCCCTTCAGGACGGCACCAATACCCATATAATAGGCCACGGAAGCTTCT 540
Db 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTGCCCGTCCACAAATCCGTTACCGCGCTGGTCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGCGGGTTACGGCATCTCCATCTCAATTCCTGGCCACAGACACACACCCCGGACG 660
Db 201 AlaValGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
QY 661 TCGCTTGATATGAATTCAATAACCTCAGCGAGTTCGTATTTTAGTCCAGCGCGGATATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGCTGGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCTGCTCTGTTATGCTTTGGATA 840

; CURRENT APPLICATION NUMBER: US/10/381,770
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 660
; TYPE: PRT
; ORGANISM: hepatitis E virus
US-10-381-770-1

Alignment Scores:

Pred. No.: 1,066-241 Length: 660
Score: 3391.00 Matches: 656
Percent Similarity: 99.55% Conservative: 1
Best Local Similarity: 99.39% Mismatches: 3
Query Match: 92.93% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-381-770-1 (1-660)

QY 1 ATGCGCCTCGCCTATTTGTTGCTGCTCTCTCATGTTTTCCTATGCTGCCGCGCA 60
DB 1 MetArgProArgProIleLeuLeuLeuLeuMetPheLeuProMetLeuProIlePro 20
QY 61 CCGCCCGGTACCGCTCTGCGCGCGCTGCTGCGCGCGCAGCGCGGTTCGCGCGTGT 120
DB 21 ProProGlyGlnProSerGlyArgArgArgGlyArgSerGlyGlyGlyGly 40
QY 121 TTCTGGGTGACCGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCAAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTrpIleHisProThrAsn 60
QY 181 CCCTTCGCCCGGATGTCACCGCTCGCGCGCGGTGACCTCGGTGTCGCCAACCCGCC 240
DB 61 ProPheAlaProAspValThrAlaAlaAlaGlyAlaGlyProArgValArgGlnProAla 80
QY 241 CGACCACTCGGCTCGCGCTGGCGTGACCAAGCGCCAGCGCCCGCGTTCCTCAGTCGT 300
DB 81 ArgProLeuGlySerAlaTrpArgAspGlnAlaGlnArgProAlaAlaSerArgArg 100
QY 301 AGACCTACACAGCTGGGCGCGCGCTAACCGCGGTGCTCCGCGCGCATGACACCGG 360
DB 101 ArgProThrThrAlaGlyAlaAlaProLeuThrAlaValAlaProAlaHisAspThrPro 120
QY 361 CCAGTGCCTGATGTCAGCTCCCGCGCGCATCTTTCGCGCGCAGTATAACCTATCAACA 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTrpAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCGCGCACCGCACTAACCTGGTTCCTTATGCGCGCCCTCT 480
DB 141 SerProLeuThrSerSerValAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCTTTTACCCCTTCAGACCGCACCAATACCAATATAATGCGCACGGAAGCTTCT 540
DB 161 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGCTCCCGTGCCCAATCCGTTACCGCGCGTGGTCCCCCAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGCGGTACGCCATCTCCATCTCATTCGCGCCACAGACCAACACCCCGACG 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATCAATCAATACCTCGAGGATGTCGTTATTTAGTCACGCCCGGCATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGATCCCAAGTAGCGCTACACTATCGTAACCAAGGCTGGCGCTCC 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGGTCTGTTATGCTTTGCATA 840

DB 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCAGTCGTAAATTCCTATACCTATACACCTATACGGTGCCTCGGCTGTG 900
DB 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCGCTTGAGCTTGAGTTTCGCAACTTACCCCGGTAAACACCAATACCGGGTC 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGTTTATCCAGCACTGTCGCCACCGCTTCGTCGCGTGCAGACGGAGCTCCGAG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACGCGCTGCTACCGCTTATGAGGACCTCTATTTTACTAGTACTAATGCT 1080
DB 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuTyrPheThrSerThrAsnGly 360
QY 1081 GTCGCTGAGATCGGCGCGGATAGCCCTCACCTGTTCACCTTGTGACACTCTGCTT 1140
DB 361 ValGlyGluIleGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTGCGCAGCAATTCATTTCGCGCTGGTGGCCAGCTGTTCTATCTCCGCTCC 1200
DB 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTCTCTCAGCAATGCGGAGCGGCTGTTAAGTTGTATATCATCTGTAGAGAACTCCTAG 1260
DB 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGATAAGGTATTGCAATCCCGCATGACATTCACCTCGGAGAACTCTGTTGTTATT 1320
DB 421 GlnAspLysGlyIleAlaIleProHisAspLeuLeuGlyGluSerArgValValIle 440
QY 1321 CAGATTATGATAACCAACATGAACAGATCGCGCAGCGCTTCTCCAGCCCATCGCGC 1380
DB 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTCTCTGCTTCGAGCTAATGATGCTGTTGGCTCTCTCACCGCTGCCGAGTAT 1440
DB 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCCACTTATGGCTCTTCGACTGGCCAGTTATGTTCTGACTCTGTGACCTTG 1500
DB 481 AspGlnSerThrTyrGlySerSerThrGlyProValTyrValSerAspSerValThrLeu 500
QY 1501 GTTAATGTTGGACCGCGCGCAGCGCTTGCCTGCGCTCGATTGACCAAGGTACA 1560
DB 501 ValAsnValAlaThrGlyAlaGlnAlaValaAlaArgSerLeuAspTrpThrLysValThr 520
QY 1561 CTTCGAGCTGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTCTTTGTCTCTGCGCG 1620
DB 521 LeuAspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTAAAGCTCTCTTCTGGAGGACGACCACTAAAGCCGGGTACCTTTATAAT 1680
DB 541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACCACTGTAGCGACCACTGTTGTCGAGATGCGCGCGCGCGCGGTGCGCT 1740
DB 561 TyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCACTAGCTGGTGTGCTGCTCCCGTCTCATTTCTGCGGTTCCTGTT 1800
DB 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATGCTTGAGGATACCTTGGACTACCTGCGCGCGC 1860
DB 601 LeuAlaProHisSerValLeuAlaLeuLeuGluAspThrMetAspTyrProAlaArgAla 620
QY 1861 CATACTTTGATGATTTCTGCCAGAGTGCAGCCCTTGGCTTTCAGGCTGCGCTTTC 1920

QY 841 CATGGCTCACTCGTAATTCCTATATAACACCTATACCGGTGCCCTCGGCTGTG 900
DB HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACCGGGTTC 960
DB AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATCCAGCACTGCTCGCACCGCTTCGTCGGTGGCGGACGACGCGCGAG 1020
DB SerArgTyrSerSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGCTGTACCCGCTTATGAAGACCTCTATTTTACTAGTACTAATGGT 1080
DB LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGly 360
QY 1081 GTCGGTGAGATCGCCGGGATAGCCCTCACCTGTTCACCTTGCTGACACTCTGCTT 1140
DB ValGlyGluValGlyArgGlyLeuAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGGCTGCGGACAGAAATTCATTCGTCGGCTGGTGGCCAGCTTCTACTCCGCTCC 1200
DB GlyGlyLeuProThrGluLeuLeuSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTCAGCAATGGCGGACCGACTCTTAAGTTGTATACATCTGTAGAAATGCTCAG 1260
DB ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGATTAAGGTATTCGAATCCCGCATGACATTCGCTCGGAGAAATCTCGTGTGTTATT 1320
DB GlnAspLysGlyValAlaLeuProHisAspLeuAspLeuGlyAspSerArgValValLe 440
QY 1321 CAGATTAATGATACCAACATGACACAGATCGCGGCGCTCTCTCAGCCCGCTCGGC 1380
DB GlnAspTyrAspAsnGlnHisGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTCTCTGTCCTTCGAGCTTAATGATGTGCTTGTGGCTCTCTCTCAGCGCTCGCGAT 1440
DB ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCAATGCTTATGCTCTTCGACTGCGGCGCACTTATGTTTCTGACTCTGTGACCTTG 1500
DB AspGlnSerThrTyrGlySerSerThrGlyProValTyrLeuSerAspSerValThrLeu 500
QY 1501 GTTAATGCTGCGCGCGCGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
DB ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 520
QY 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCACTACTCGAAGACCTTCTTGTCTGCTGCG 1620
DB LeuAspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuPro 540
QY 1621 CTCGCGGTGAGCTCTCTTCTGGGAGCGAGCGCACTAAAGCGCGGTACCTTTAAT 1680
DB LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACCACTGCTAGCGACCACTCTCTCGAGATCGCGGCGCGCGCGCGCTGCT 1740
DB TyrAsnThrThrAlaSerAspGlnIleLeuLeuGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCACTAGCTGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB IleSerThrTyrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTCCTTGAGGATACCTTTGAGTACCTCGCGCGCGCC 1860
DB LeuAlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAla 620
QY 1861 CATACTTTTGATATTCGCGCGAGTGGCGCGCGCTTGGCTTACGGGTGCGCTTTC 1920
DB HisThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTGCTGAGCTTCAGCGCTTAAGATGAAGTGGGTAAACTCGGAGTTG 1980

DB 641 GlnSerThrValAlaGluLeuGlnArgLeuLysValGlyLysThrArgGluLeu 660
RESULT 7
US-10-165-868-20
; Sequence 20, Application US/10165868
; Publication No. US20030143241A1
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; Bradley, Daniel W.
; Twu, Jr-Shin
; Purdy, Michael A.
; Tam, Albert W.
; Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/165,868
; FILING DATE: 06-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 9
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-10-165-868-20
Alignment Scores: 3.45e-227 Length: 659
Pred. No.: 3194.50 Matches: 614
Score: 96.21% Conservative: 21
Best Local Similarity: 93.03% Mismatches: 24
Query Match: 87.54% Indels: 1
DB: 14 Gaps: 1

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-165-868-20 (1-659)

QY 1 ATGCGCCCTCGCCCTATTGTTGCTGCTCCTCATGTTTTCCTATGTCGCCCGGCCA 60
Db 1 MetArgProArgProLeuLeuLeuPheLeuLeuPheLeuLeuMetLeuProAlaPro 20
QY 61 CGCGCCGGTACGCGCTCTGCGCGCGCTGCTGCGCGCGCGCGCGCGGTTCGCGCGTGT 120
Db 21 ProThrGlyInProSerGlyArgArgGlyArgGlyArgGlyArgGlyArgGly 40
QY 121 TTCTGGGGTACCGGGTGAATCTCAGCCCTTCGCAATCCCTATATCATCAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCATGTCACCGCTCGCGCGCGGGGTGACCTGCTGTCGCCAACCGGC 240
Db 61 ProPheAlaProAspValAlaAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGACCACTCGGCTCGGCTTGGCGTACACGAGCCCGCGCGCGCGCGCTTGCCTCAGTGT 300
Db 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
QY 301 AGACCTACACAGTGGCGCGCGCTAACCGCGTGTCCCGCCCATGACACCCCG 360
Db 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCAGTGCCTGATGTCGATCCCGCGCGCATCTTGGCGCGCGAGTATACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGTCGCCACCGGCACCTAACCTGTTCTTTATGCGCCCTCTT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGCTTTTACCTCTCAGACCGCACCAATACCATATATATGCGCACGGAAGTCT 540
Db 161 AsnProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGTTCCCGTCCGACCAATCCGTACCGCGCGCTGTCGCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GTGTGCGCGGTAGCCATCTCCATCTCATTCTGCGCACAGACACACACACCCCGACG 660
Db 201 AlaValGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATGAATCAATACCTCGACGAGTGTGCTATTTTATGTCAGCCCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTGTGATCCCAAGTCAGCGCCTACACTATCGTAACCAAGGCTGCGCTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTCCAGACCTCTGGGGTGGCTGAGGAGGCTACTCTGCTGTTGTTATGCTTTGCGATA 840
Db 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCAGTCGTAATTCCTATCTAATAACCCCTATACCGGTGCGCTCGGGCTGTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCGCTTGAGCTTGAGTTTCGACCTTACCCCGGTAAACCAATACCGGGTTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTATTCAGCACTGCTCGCCACCGCCCTTCGTCGGGTGCGGACGGAGCTCCGAG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisSer---AlaArgGlyAlaAspGlyThrAlaGlu 339
QY 1021 CTCACACACGGCTGCTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGGT 1080
|||||

Db 340 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGly 359
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCCTGTTCAACCTTGCCTGACACTGCTT 1140
Db 360 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeu 379
QY 1141 GCGCGCTCGCGCACAGAATTGATTTCGTCGGTGTGGCGCAGCTGTTCTACTCCCGTCC 1200
Db 380 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 399
QY 1201 GTTGTCTCAGCAATGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAAATGCTCAG 1260
Db 400 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 419
QY 1261 CAGGATAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAAATCTGCTGGTTATT 1320
Db 420 GlnAspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValIle 439
QY 1321 CAGGATATTGATAACCAACATCAACAGATCGCGCGAGCGCTTCTCCAGCCCATCGCGC 1380
Db 440 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 459
QY 1381 CTTTCTCTGCTTCGAGCTAATGATGTGCTTTCGCTCTCTCTCACCGCTCCGAGTAT 1440
Db 460 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 479
QY 1441 GACCACTCCACTATGCTCTTCGACTGCGCGCAGTTATGTTTCTGACTCTGCTGACCTG 1500
Db 480 AspGlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeu 499
QY 1501 GTTAATGTTGCGACCGCGCGCAGCGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db 500 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 519
QY 1561 CTTGACGCTGCGCCCTCTCCACCATCCAGCAGTACTGAGAGACCTTCTTTCCTCTGCGC 1620
Db 520 LeuAspGlyArgProLeuProThrValGluGlnTyrSerLysThrPheValLeuPro 539
QY 1621 CTCGCGGTGAGCTCTCTTCTGGGAGGAGCAGCACAACTAAAGCCGGGTACCTTATAT 1680
Db 540 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 559
QY 1681 TATAACCACTGCTAGCGACCACTGCTGTGCGAATGCGCGCGGCGCACCGGCTCGCT 1740
Db 560 TyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAla 579
QY 1741 ATTTCCACTTACCACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 580 IleSerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaVal 599
QY 1801 TTAGCCCCCACTCTGCGTAGCTGCTTGGAGGATACCTTGCAGTACCTGCTGCTGCTGCTGCT 1860
Db 600 LeuAlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAla 619
QY 1861 CATACTTTGATTTCTGCGCAGAGTCCCGCCCTTGGCTTTCAGGCTGCTGCTTTC 1920
Db 620 HisThrPheAspAspPheCysProGluCysArgAlaLeuLeuGlyLeuGlnGlyCysAlaPhe 639
QY 1921 CAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 640 GlnSerThrValAlaGluLeuGlnArgLeuLysValLysValLysThrArgGluLeu 659

RESULT 8

US-10-239-090A-57

; Sequence 57, Application US/10239090A

; Publication No. US20040101820A1

; GENERAL INFORMATION:

; APPLICANT: KABUSHIKI KAISHA TOSHIBA

; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co

; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,

; TITLE OF INVENTION: detecting hepatitis E virus using the same

; FILE REFERENCE: 02S0741P

; CURRENT APPLICATION NUMBER: US/10/239,090A


```
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 57
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Hepatitis E Virus JAK-Sai (ORF2)
US-10-239-090A-57

Alignment Scores:
Pred. No.: 4,84e-226 Length: 674
Score: 3179.00 Matches: 607
Percent Similarity: 95.90% Conservative: 25
Best Local Similarity: 92.11% Mismatches: 27
Query Match: 87.12% Indels: 0
DB: 16 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-239-090A-57 (1-674)

QY 1 ATGCGCCTCGGCTATTGTTGCTGCTCCTCATGTTTTTGGCTATGCTGCGCGGCCA 60
DB 15 MetArgSerArgAlaLeuPheLeuPheLeuPheValPheLeuProMetLeuProAlaPro 34
QY 61 CCGCCCGGTCAGCGCTCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGT 120
DB 35 ProAlaGlyGlnProSerGlyArgArgGlyArgSerGlyGlyAlaGlyGlyGly 54
QY 121 TTCTGGGGTGACCGGGTGTATCTCAGCCCTTCCCAATCCCTATATTCATCCAAAC 180
DB 55 PheTrpGlyAspArgValAspSerGlnProPheAlaLeuProTyrIleHisProThrAsn 74
QY 181 CCCTTCGCCCGCGATGTCACCGCTGCGCGCGCGCGCTGCGCGCTGTTGCGCACCGCC 240
DB 75 ProPheAlaSerAspIleProThrAlaAlaGlyAlaGlyAlaArgProArgGlnProAla 94
QY 241 CGACCACTCGGCTCGGCTTGCGGTGACACGCGCGCGCGCGCGCGCGCTGCTCAGTCGT 300
DB 95 ArgProLeuGlySerAlaTrpArgAspGlnSerGlnArgProThrAlaSerAlaArgArg 114
QY 301 AGACCTACCAAGCTGGGGCGCGCGCTAACCGCGGTGCGTCTCGGCGCGCGTACACCCG 360
DB 115 ArgProAlaProAlaGlyAlaAlaProLeuThrAlaValAlaProAlaProAspThrAla 134
QY 361 CCAGTGCCTGATGTCGACTCCCGCGCGCGCATCTGCGCGCGCGCATACCTATCAACA 420
DB 135 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 154
QY 421 TCTCCCTTACCTCTCCGTTGCGCGCGCGCATACCTGCTTCTTTATGCGCGCGCTCTT 480
DB 155 SerProLeuThrSerThrIleAlaThrGlyThrAsnLeuValLeuTyrAlaAlaProLeu 174
QY 481 AGTCCGCTTTTACCTCTCAGGCGCGCGCATACCTATACCTATATATGCGCGCGAGCTTCT 540
DB 175 SerProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 194
QY 541 AATTTATGCCAGTACCGGGTTCGCGTGCACCAATCCGTTACCGCGCGCGTGGTCCCAAT 600
DB 195 AsnTyrAlaGlnTyrArgValValArgAlaThrIleArgTyrArgProLeuValProAsn 214
QY 601 GCTGTGCGCGGTTACGCGCATCTCCATCTCATCTGCGCGCGCGCGCGCGCGCGCGCG 660
DB 215 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 234
QY 661 TCGTTGATATGAATCAATAACCTCAGCGGATGTTCTGATTTTATGTCAGCGCGGATA 720
DB 235 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 254
QY 721 GCCTCTCAGCTGTGATCCCAAGTGAGCGCTACACTATCGTACCAAGCTGCGCGTCC 780
DB 255 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTTrpArgSer 274
QY 781 GTCGAGACCTCTGCGGGTGGCTGAGGAGGCGCTACCTCTGCTCTTGTATGCTTTGCATA 840
```

```
DB 275 ValGlnThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 294
QY 841 CATGGCTCAGCTCGTAATTCCTATACATACACCTATACCGGTGCGCTCGGGCTGTG 900
DB 295 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 314
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTACCAACCAATACCGGGT 960
DB 315 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 334
QY 961 TCCCGTTATTCCAGCACTGCTCGCCACCGCTTCTGTCGCGTGGCGGCGGACTGCCGAG 1020
DB 335 SerArgTyrSerSerSerAlaArgHisIlyLeuArgGlyProAspGlyThrAlaGlu 354
QY 1021 CTACACACCGGCTGCTACCGCTTATGAGAGACCTCTATTTTACTAGTACTAATGGT 1080
DB 355 LeuThrThrThrAlaAlaThrArgPheMetIlyAspLeuHisPheThrGlyThrAsnGly 374
QY 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTGTTCAACCTTGTGCTGACACCTCGCTT 1140
DB 375 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 394
QY 1141 GCGCGCTCGCGCAGAAATTGATTTCGCTGCGCTGCGCGCTGCTGCTTCTACTCCCGTCCC 1200
DB 395 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 414
QY 1201 GTTGCTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTATACATCTGTAGAGAATGCTCAG 1260
DB 415 ValValSerAlaAsnGlyGluProThrValIlyLeuTyrThrSerValGluAsnAlaGln 434
QY 1261 CAGGATAAGGGTATTGCAATCCGCGATGACCTCGGAGAACTCTCGTGTGGTTATT 1320
DB 435 GlnAspIlyGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValIle 454
QY 1321 CAGGATTATGATAACCAACATGAACAAGATCGGCGCGCGCTTCTCCAGCGCCCACTCGGC 1380
DB 455 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 474
QY 1381 CTTTCTCTGCTTCTGAGCTAATGATGCTGTTGGCTCTCTCTCACCGCTGCGGAGTAT 1440
DB 475 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 494
QY 1441 GACCAGTCCACTTATGCTCTCGACTGCGCGCGCGCTTATGCTGACTCTGCTGACCTTG 1500
DB 495 AspGlnThrThrTyrGlySerSerThrAsnProMetTyrValSerAspThrValThrPhe 514
QY 1501 GTTAATGTTCCGACCGCGCGCGCGCTTGGCGCGCTCGCTCGATTGCGACCAAGTCCACA 1560
DB 515 ValAsnValAlaThrGlyAlaGlnGlyValSerArgSerLeuAspTrpSerLysValThr 534
QY 1561 CTTGACCGGTCGCGCGCTCTCCACCATCCAGAGTATCTCGAAGACCTTCTTTGCTCGCGC 1620
DB 535 LeuAspGlyArgProLeuThrThrIleGlnGlnTyrSerLysThrPhePheValLeuPro 554
QY 1621 CTCGCGGTAAAGCTCTCTTCTGGGAGGAGGACACATAAGCCGGGTACCTTATAAT 1680
DB 555 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 574
QY 1681 TATAACCACTGCTAGCGACCAACTGTTGTCGAGAATCCCGCGCGCGCGCGCTCGCT 1740
DB 575 TyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValCys 594
QY 1741 ATTTCCACTTACACCTAGCTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB 595 IleSerThrTyrThrThrAsnLeuGlySerGlyProValSerIleSerAlaValGlyVal 614
QY 1801 TTAGCCCCCAGCTCTGCGCTAGCATTCCTTGGAGATACCTTGGACTACCTGCGCGCGC 1860
DB 615 LeuAlaProHisSerAlaLeuAlaLeuGluAspThrValAspTyrProAlaArgAla 634
QY 1861 CATACTTTTTCATGATTTCTGCCAGAGTGCAGCGCGCGCTGCTGCTGCTGCTGCTGCTTTC 1920
```

Db 635 HisThrPheAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 654
Qy 1921 CAGTCTACTGTCGCTGAGCTTCAGCGCTTAAGATGAAGTGGGTAAACTCGGAG 1977
Db 655 GlnSerThrValAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu 673
RESULT 9
US-10-239-090A-51
; Sequence 51, Application US/10239090A
; Publication No. US20040101820A1
; GENERAL INFORMATION:
; APPLICANT: KABUSHIKI KAISHA TOSHIBA
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co
; TITLE OF INVENTION: from Japanese chip including the same, kit including the same,
; TITLE OF INVENTION: detecting hepatitis E virus using the same
; FILE REFERENCE: 0280741P
; CURRENT APPLICATION NUMBER: US/10/239,090A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 51
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus JKN-Sap (ORF2)
US-10-239-090A-51
Alignment Scores:
Pred. No.: 8,03e-226 Length: 660
Score: 3176.00 Matches: 607
Percent Similarity: 96.05% Conservative: 26
Best Local Similarity: 92.11% Mismatches: 26
Query Match: 87.04% Indels: 0
Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-239-090A-51 (1-660)
Qy 1 ATGGCCCTCGCGCTATTGTTGCTGCTCCCTCANGTTTTCCTATGCTGCGCGCCA 60
Db 1 MetArgProArgAlaValLeuLeuLeuPheLeuValLeuProMetLeuProAlaPro 20
Qy 61 CGCGCGCTGACCGCTGCGCGCGCTGCTGGCGCGCGCGCGCGCTGCGCGGTG 120
Db 21 ProAlaGlyGlnProSerGlyArgArgGlyArgSerGlyGlyAlaGlyGly 40
Qy 121 TTCTGGGTGACCGGTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaLeuProTyrlleHisProThrAsn 60
Qy 181 CCTTCGCGCGCGATGTCAACCGCTGCGCGCGCGCTGGACCTCGTGTTCGCCAACCCGCGC 240
Db 61 ProPheAlaAlaAspValValSerGlnProGlyAlaGlyThrArgProArgGlnProPro 80
Qy 241 GCACACTCGGCTCGCTGCGCTGACAGGCGCGCGCGCGCGCTGCGCTGCGCTGCT 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnSerGlnArgProSerAlaAlaProArg 100
Qy 301 AGACCTACACAGTGGCGCGCGCTAACCGCGCTGCTGCGCGCGCGCGCGCGCGCG 360
Db 101 ArgSerAlaProAlaGlyAlaAlaProLeuThrAlaValSerProAlaProAspThrAla 120
Qy 361 CCAGTGCCTGATGTCGACTCCCGCGCGCGCATCTTGGCGCGCGCATTAACCTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaLeuLeuArgGlnGlnTyrlleHisProThr 140
Qy 421 TCTCCGCTTACTCTTCCGTGCGCACCGGCTATACCTGGTCTTTATGCTGCTGCTGCT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrlleAlaProLeu 160
Qy 481 AGTCCGCTTTTACCGCTTCAGGACCGCACCATATACATATATATATATATATATAT 540
Db 161 AsnProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180

Qy 541 AATTATGCCAGTACCGGTTGCCCGTCCCAATCCGTTACCGCCCGCTGCTGCCCAAT 600
Db 181 AsnTyrlleGlnTyrlleArgValValAlaThrIleArgTyrlleArgProLeuValProAsn 200
Qy 601 GCTGTCGCGGTTCAGCCATCTCCATCTCATTCTGGCCACAGCACACACACCCGCGAG 660
Db 201 AlaValGlyGlyTyrlleAlaIleSerPheTrpProGlnThrThrThrProThr 220
Qy 661 TCCGTTGATATGAATTCATTAACCTCGACGAGTGTTCGTTATTTAGTCCAGCCCGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
Qy 721 GCCTCTGAGCTTGATCCCAAGTCCGAGCGCTACACTATCGTAACCAAGGCTGGCGTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrlleArgAsnGlnGlyTrpArgSer 260
Qy 781 GTCGAGACCTCTGGGTGCTGAGGAGGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 261 ValGluThrThrGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
Qy 841 CATGCTCACCTCGTAAATTCCTATATAATACACCTATACCGGTCCCTCGGCTGTG 900
Db 281 HisGlySerProValAsnSerTyrlleAsnThrProTyrlleGlyAlaLeuGlyLeuLeu 300
Qy 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGTC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
Qy 961 TCCGCTTATTCAGCACTGCTGCGCACCGCTTCTGCGGTGCGGAGCGGAGCTCCGAG 1020
Db 321 SerArgTyrlleThrSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
Qy 1021 CTCACCAACCGCTGCTACCGCTTATGAAGGACCTCTATTTTACTAGTACTAATGCT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyThrAsnGly 360
Qy 1081 GTCGCTGAGATGCGCGCGGATAGCCCTCACCTGTTCAACCTTCTGACACTCTGCT 1140
Db 361 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
Qy 1141 GCGCGCTGCGGACAGAAATGATTCGCGGTGCTGCGGTGCGGAGCTTCTACTCCCGTCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrlleSerArgPro 400
Qy 1201 GTTGCTCAGCCCAATGCGGAGCGGCTGTTAGTTGTATACATCTGTAGAGAACTCTAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrlleSerValGluAsnAlaGln 420
Qy 1261 CAGGATAAGGTPATTGCAATCCCGCATGACATGACCTCGGAGAACTCTGCTGCTGCTAT 1320
Db 421 GlnAspLysGlyIleThrIleProHisAspIleAspLeuGlyAspSerArgValIle 440
Qy 1321 CAGGATATGATAACCAATGAAACAGATCGCGCGAGCTTCTCCAGCCCGCATCGCG 1380
Db 441 GlnAspTyrlleAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
Qy 1381 CTTTCTCTGCTCTCGAGCTAATGATGCTTTCGCTCTCTCTCACCGCTCCCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGlyTyrlle 480
Qy 1441 GACCACTCCACTATGCTCTTCGACTGCGCCAGTTTATGTTTCTGACTCTGCTGCTGCT 1500
Db 481 AspGlnThrThrTyrlleSerSerThrAsnProMetTyrlleValSerAspThrValThrLeu 500
Qy 1501 GTTATGTTGCGACCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 520
Qy 1561 CTTGACGCTGCGCGCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTCTGCTGCTGCG 1620
Db 521 LeuAspGlyArgProLeuThrThrIleGlnTyrlleSerTyrllePheTyrlleValLeuPro 540
Qy 1621 CTCGCGGTGAGCTCTCTTCTGCGGAGGCGAGCACACATAAGCCGGGTACCTTATAT 1680

```
Db 541 LeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACCACTGCTAGGACCACTGCTTCGAGATGCGCGGCACCGGTGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnLeuLeuGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACACTAGCTGGGTGCTGCTGCCGCTCCATTCTCTGCGGTTCGCGTT 1800
Db 581 IleSerThrThrThrSerLeuGlyAlaGlyProThrSerIleSerAlaValGlyVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCACTGCTGAGGATACCTTGGATACCTCTGCCCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaValLeuGluAspThrValAspTyrProAlaArgAla 620
QY 1861 CATACTTTTATGATTTCTGCCGAGTGGCGCCCGCTTGGCCCTCAGGGTGGCGTTTC 1920
Db 621 HisThrPheAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGCTGAGCTTCAGCGCTTAAGATGAAGTGGGTAAACCTCGGGAG 1977
Db 641 GlnSerThrIleAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu 659

RESULT 10
US-10-239-090A-53
; Sequence 53, Application US/10239090A
; Publication No. US20040101820A1
; GENERAL INFORMATION:
; APPLICANT: KABUSHIKI KATSHA TOSHIBA
; TITLE OF INVENTION: Polynucleotide probe and primer derived from hepatitis E virus co
; TITLE OF INVENTION: from Japanese, chip including the same, kit including the same,
; TITLE OF INVENTION: detecting hepatitis E virus using the same
; FILE REFERENCE: 02S0741P
; CURRENT APPLICATION NUMBER: US/10/239,090A
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: JP2001-191837
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 57
; SEQ ID NO 53
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus JMY-Haw (ORF2)
US-10-239-090A-53

Alignment Scores:
Pred. No.: 1,13e-225 Length: 660
Score: 3174.00 Matches: 607
Percent Similarity: 95.90% Conservative: 25
Best Local Similarity: 92.11% Mismatches: 27
Query Match: 86.98% Indels: 0
DB: 16 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1993) x US-10-239-090A-53 (1-660)
QY 1 ATGCGCCTCGGCCTATTTTGTCTGCTCTCTCATGTTTTTGTCTGCTGCGCGCCA 60
Db 1 MetArgProArgAlaValLeuLeuLeuPheLeuValLeuLeuProMetLeuProAlaPro 20
QY 61 CCGCCCGGTGACCGCTGCGCGCGGTGCTGCGCGCGCGCGCGCGGTTCGCGGTGGT 120
Db 21 ProAlaGlyGlnProSerGlyArgArgGlyArgSerGlyThrGlyGly 40
QY 121 TTCTGGGGTGACCGGGTGTATTCACGCTTCGCAATCCCTATATTCATCCAAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaLeuProTyrIleHisProThrAsn 60
QY 181 CCCTTCGCCCGCATGTGACCGCTGCGCGCGCGGTGCGACCTGCTGTTTCGCAACCGCC 240
Db 61 ProPheAlaAlaAspValValSerGlnProGlyAlaGlyAlaArgProArgGlnProPro 80
QY 241 CGACCACTCGGCTCGGTGCGGTGACACGAGCCCGCGCGCGCGGTGCTACGTCGT 300
Db 81 ArgProLeuGlySerAlaTrpArgAspGlnSerGlnArgProSerAlaAlaProArgArg 100
```

```
QY 301 AGACCTACCAGCTGGCGCGCGCGCTAACCGGGTTCGCTCGGCGCCCATGACACCCCG 360
Db 101 ArgSerAlaProAlaGlyAlaAlaProLeuThrAlaValSerProAlaProAspThrAla 120
QY 361 CCAGTGGCTGATGTGCGACTTCCCGCGCGGCATCTTTCGCGCGCGCAGTATAACTATCAACA 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCTCCCTTACCTCTTCGCGCGCACCGGACACTAACCTGGTTCTTTATGCGCGCGCTCTT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AGTCGGCTTTTACCCCTTCAGGACGGCACCAATACCATTAATGCGCCACCGGAGCTTCT 540
Db 161 AsnProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCCAGTACCGGGTTGCCGTCGCCACAAATCCGTTTACCAGCCCGCTGGTCCCCAAT 600
Db 181 AsnTyrAlaGlnTyrArgValValArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCTGTCGGCGGTTCAGCCATCTCCATCTCATTCCTGGCCACAGACACACACCCCGCAGC 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCCGTTGATATGAATTCATAAATCCTCAGCGATGTTTCGTATTTTAGTCCAGCCCGGAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCCTCTGAGCTTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGTGGCGTCC 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
QY 781 GTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTACCTCTGCTGCTTGTATGCTTTGATG 840
Db 261 ValGluThrThrGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCACTCGTAAATTCCTACTACTAATACACCTATACCGGTGCGCTGGCGCTGTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTCC 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGGTTATTCAGCACTGCTCGCCACCGCTTCGTCGCGGTGCGGAGCGGACTGCGCAG 1020
Db 321 SerArgTyrThrSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCAACGCTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGCT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyThrAsnGly 360
QY 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTTCAACCTTGTGACACTCTGCTT 1140
Db 361 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GGGCGCTCGCGCAGAAATTGATTTGTCGCGGTGCGGCGCAGCTGTTCTACTCCCGTCCC 1200
Db 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTGTCTTCAGCAATGCGGAGCGCGCTGTTTAAAGTGTATACATCTGTAGAGATGCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGGTATTTGCAATCCCGCATGACATTGACCTCGGAGAACTCTCGTGGTTATT 1320
Db 421 GlnAspLysGlyIleThrIleProHisAspIleAspLeuGlyAspSerArgValValIle 440
QY 1321 CAGGATTATGATTAACCAACATGAACAGATCGGCGCAGCGCTTCTCCAGCCCCCATCGCG 1380
Db 441 GlnAspTyrAspAsnGlnHisGluAsnArgProThrProSerProAlaProSerArg 460
```


QY 721 GCCTCTGAGCTTGATCCCAAGTCAGCGCCTACACTATCGTAACCAAGGCTGGCGCTCC 780
Db 241 AlaserGluLeuValleProSerGluargGluHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 CTCGAGACCTCTGGGGTGGCTCAGGAGGAGCTACCTCTGGTCTTGTATGCTTTGCATA 840
Db 261 ValGluThrThrGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuGlyCysile 280
QY 841 CATGGCTCACTCGTAATTCCTACTAATACACCTATACCGGTGCGCTCGGCTGCTTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGGCTTACCTTGGTTCGCAACCTTACCCCGGTAAACCAATACGCGGCTC 960
Db 301 AspPheAlaLeuGluLeuGluPheargAsnLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961 TCCGGTATTCCAGACACTGCTCGCACCGCTTCCTCGGTGCGGAGCGGAGTCCGAG 1020
Db 321 SerArgTyrThrSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
QY 1021 CTCACCAACCGGCTGCTACCGCTTATGAGGACCTTATTTACTAGTACTAATGCT 1080
Db 341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyThrAsnGly 360
QY 1081 GTCGGTGAGATCGCGCGGGATAGCCCTCACCTGTTCAACCTTGCTGACACTCTGCTT 1140
Db 361 ValGlyGluValGlyArgGlyleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141 GCGCGCTCGCGACAGAAATTGATTCGTCGGCTGTCGCGAGCTTCTTACTCCGCTCC 1200
Db 381 GlyGlyLeuProThrGluLeuileSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201 GTTCTCTCAGCAATGGCGCGGCTGTTAAGTTGTATACATCTGAGAGACTCTCAG 1260
Db 401 ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261 CAGGATAAGGGTATTGCAATCCCGCATCGCATTCACCTCGGAGAACTCTGCTGCTTATT 1320
Db 421 GlnAspLysGlyleThrleProHleAspIleAspLeuGlyAspSerArgValValle 440
QY 1321 CAGGATTTATGATAACAATGAAACAAGATCGCGCGAGCTTCTCCAGCGCCATCGCGC 1380
Db 441 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381 CCTTCTCTGCTTCGAGCTAATGATGCTTTCGCTCTCTCACCGCTGCCAGAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuThrLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCACTCCACTATGCTCTTCGACTCGCGCCAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 481 ***GlnThrThrTyrGlySerSerThrAsnProMetTyrValSerAspThrValThrLeu 500
QY 1501 GTTAATGTTGCGACCGCGCGAGCGCTTGGCGGCTGCTCGATGAGACCAAGTCA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 520
QY 1561 CTGACGCTGCGCCCTCTCCACCATCCAGAGTACTCGAGACCTTCTTGTCTCTGCGG 1620
Db 521 LeuAspGlyArgProLeuThrThrleGlnGlnTyrSerLysLysPheTyrValLeuPro 540
QY 1621 CTCGCGGTAACTCTCTTCTGGGAGGCGAGCAACAACCTAAAGCGGGTACCTTTAAT 1680
Db 541 Leu***GlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATAACACACTGCTAGACCAACTGCTTGTGAGAAATGCGCGCGGACCGGCTCGCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnleLeuileGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACACTAGCTGGTGTGCTGCCGCTCTCCATTTCTGCGGTTCCGCTT 1800
Db 581 lleSerThrTyrThrThrSerLeuGlyAlaGlyProThrSer***SerAlaValGlyVal 600

QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGTACATCTTGAGTACCCTCCCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaValLeuGluAspThrValAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGTATGATTTCTGCCAGAGTGGCGCCCTTGGCTTCAGGCTTCGCGCTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTGAGCTTACGCGCTTAAGATGAAGTGGGTAAACTCGGGAG 1977
Db 641 GlnSerThrIleAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu 659
RESULT 13
US-10-319-745-92
; Sequence 92, Application US/10319745
; Publication No. US20030211467A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushanwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/10/319,745
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: US/09/468,147A
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/173,141
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/061,199
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 92
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 481
; OTHER INFORMATION: Xaa = Unknown or Other at position 542
; OTHER INFORMATION: Xaa = Unknown or Other at position 595
US-10-319-745-92
Alignment Scores:
Pred. No.: 1.22e-224 Length: 660
Score: 3160.00 Matches: 604
Percent Similarity: 95.75% Conservative: 27
Best Local Similarity: 91.65% Mismatches: 28
Query Match: 86.60% Indels: 0
DB: 12 Gaps: 0
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-319-745-92 (1-660)
QY 1 ATGCGCCCTCGCGCTATTGTTGTTGCTGCTCTCATGTTTTCCTATGTTGCTCCCGCA 60
Db 1 MetArgProArgAlaValLeuLeuLeuPheLeuMetPheLeuProMetLeuProAlaPro 20
QY 61 CGCGCGGTCAGCGCTCTGCGCGCGCTGCTGGCGCGCGAGCGCGGTTCCGCGCGTGT 120
Db 21 ProAlaGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlyAlaGlyGlyGly 40
QY 121 TTCTGGGTGACCGGTTGATTCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
Db 41 PheTrpSerAspArgValAspSerGlnProPheAlaLeuProTyrIleHisProThrAsn 60
QY 181 CCTTCGCGCCCGGATGTACCGCTGCGCGCGCGGCTGGACCTCGTGTTCGCCAACCCGCC 240
Db 61 PropheAlaAlaAspValValSerGlnProGlyAlaGlyThrArgProArgGlnProPro 80
QY 241 CGACCACTCGGCTCGCTTGGGTGACCGGCCCGCGCGCGCGCTTCCTCAGTCGT 300

|||||
81 ArgProLeuGlySerAlaThrArgAspGlnSerLysArgProSerValAlaProArgArg 100
|||||
301 AGACCTTACCAGCTGGGGCGCGCGCTAACCGGGTGCCTCGCGCCCATGACACCCCG 360
|||||
101 ArgSerThrProAlaGlyAlaAlaProLeuThrAlaIleSerProAlaProAspThrAla 120
|||||
361 CCAGTGCCTGATGTCGACTCCCGGGCGCCATCTTGGCGCGCAGTATAACCTATCAACA 420
|||||
121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
|||||
421 TCTCCCTTACTCTTCCGTGGCCACCGGCACCTAACCTGGTCTTTATGCGCCCTCTT 480
|||||
141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
|||||
481 AGTCGGCTTTTACCCTTCAGAGCGGACCAATACCCATATTAATGGCCACGAGCTTCT 540
:::|||||
161 AsnProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
|||||
541 AATTATGCCAGTACCGGGTTGCCGTCGCCAATCCGTTACCGCGCGCTGCCCAAT 600
|||||
181 AsnTyrAlaGlnTyrArgValValArgAlaThrIleArgTyrArgProLeuValProAsn 200
|||||
601 GCTGTCGGCGGTACGCGCATCTCCATCTCTCATTTCTGGCCACAGACACACACCCCGACG 660
|||||
201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
|||||
661 TCGGTTGATGATGAATTCATTAACCTCGAGCGATGTTCTGATTTTAGTCCAGCCCGGATA 720
|||||
221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
|||||
721 GCCTCTGAGCTGTGATCCCAAGTGAGCGCTACACTATCGTAACCAAGGTGGCGCTCC 780
|||||
241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
|||||
781 GTCGAGACCTCGGGGTGGCTGAGGAGGAGGCTACCTCTGCTTGTGTTATGTTTGATA 840
|||||
261 ValGluThrThrGlyValAlaGluGluAlaThrSerGlyLeuValIleLeuCysIle 280
|||||
841 CATGGCTCACTCGTAATTCCTACTAATACACCTATACGGTGCCTCGGGCTGTG 900
|||||
281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
|||||
901 GACTTTGCCCTGAGTGGAGTTTCGCAACCTTACCCCGGTAAACCAATACGCGGTC 960
|||||
301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
|||||
961 TCCCGTTATTCCAGCACTGCTCGCCACCGCCTTCGTGCGGTGCGGAGCTGCCGAG 1020
|||||
321 SerArgTyrThrSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
|||||
1021 CTCACCACGCGTGTCTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGGT 1080
|||||
341 LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyThrAsnGly 360
|||||
1081 GTCGCTGAGTACGCGCGGGATAGCCCTCACCCCTGTTCAACCTTGTCGACACTCTGCTT 1140
|||||
361 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
|||||
1141 GCGCGCTTCGCGCACAGAATTGATTTTCGTGCGGTGGTGGCCAGCTCTTCTACTCCCGTCCC 1200
|||||
381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
|||||
1201 GTTGTCTCAGCAATGGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAGATCTCAG 1260
|||||
401 ValValSerAlaAsnGlyGluProThrProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
|||||
1261 CAGGTAAGGTTATTCGATCCCGCATGACATTGACCTCGGAGATCTCGTGGTATT 1320
|||||
421 GlnAspLysGlyIleThrIleProHisPheIleAspLeuGlyAspSerArgValValIle 440
|||||
1321 CAGGATTATGATACCAACATCAACAAGATCGCGCAGCCCTTCTCCAGCCCATCGCGC 1380
|||||

Db 441 GlnAspTyrAspAsnGlnHisGluAlaAspArgProThrProSerProAlaProSerArg 460
QY CTTTCTCTGCTTCGAGCTAAATGATGCTGTTGGCTCTCTCACCGCTGCCGAGTAT 1440
Db 461 ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441 GACCAGTCCACTTATGGCTCTTCGACTGGCCAGCTTTATGTTTCTGACTCTGTGACCTTG 1500
Db 481 ***GlnThrThrTyrGlySerSerThrAsnProMetTyrValSerAspThrValThrLeu 500
QY 1501 GTTAAATGTTCGACGCGCGCGACCGCTTCCCGGTGCTCGATTCGATGGACCAAGTCA 1560
Db 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 520
QY 1561 CTTGACCGTCCGCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCTTTGTCTGCTGCG 1620
Db 521 LeuAspGlyArgProLeuThrThrIleGlnGlnTyrSerLysLysPheTyrValLeuPro 540
QY 1621 CTCGCGCTTAAGCTCTCTTTCTGGGAGCGACGACACTAAAGCGGGTACCTTTATAT 1680
Db 541 Leu**GlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681 TATTAACCACTGCTAGCAGCAACTGCTTCGAGAAATCGCGCGGACCGGGTCCCT 1740
Db 561 TyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAla 580
QY 1741 ATTTCCACTTACACCACTAGCTGGGTGCTGCTCCGCTCCATTTCTGCGGTTGCCCTT 1800
Db 581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProThrSer***SerAlaValGlyVal 600
QY 1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTGAGGATACCTTGGACTACCTGCCCGCGCC 1860
Db 601 LeuAlaProHisSerAlaLeuAlaValLeuGluAspThrValAspTyrProAlaArgAla 620
QY 1861 CATACTTTTGATGATTTCTGCCAGAGTCCGCCCGCTTGGCCTTCAGGCTGCGCTTC 1920
Db 621 HisThrPheAspAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921 CAGTCTACTGTCGCTGAGCTTCAGCGCTTAAGATGAAGTGGTAAAACTCGGAG 1977
Db 641 GlnSerThrIleAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu 659

RESULT 14

US-09-468-147-167
; Sequence 167, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.PI
; CURRENT APPLICATION NUMBER: US/09/468,147A
; EARLIER FILING DATE: 1999-12-21
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 167
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 84
; OTHER INFORMATION: Xaa = Unknown or Other at position 230
; OTHER INFORMATION: Xaa = Unknown or Other at position 447
US-09-468-147-167

Alignment Scores:
Pred. No.: 5,63e-224 Length: 660
Score: 3151.00 Matches: 603
Percent Similarity: 95.45% Conservative: 26
Best Local Similarity: 91.50% Mismatches: 30
Query Match: 86.35% Indels: 0
DB: 10 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-09-468-147-167 (1-660)

```
QY 1  ATGCGCCCTCGCGCTATTTTGTGCTCCTCATGTTTGTGCTATGCTGCGCGGCCA 60
DB 1  MetArgProArgAlaValLeuLeuLeuPheValLeuLeuProMetLeuProAlaPro 20

QY 61  CGCGCGGTGACCGCTGCTGCGCGCGTGTGGCGCGCGAGCGGGTTCGGCGGTGT 120
DB 21  ProAlaGlyGlnProSerGlyArgArgGlyArgSerGlyGlyAlaGlyGlyGly 40

QY 121  TTCTGGGGTGACCGGTGATCTCAGCCCTTCGCAATCCCTATATTCATCCAAACAAC 180
DB 41  PheTrpGlyAspArgValAspSerGlnProPheAlaLeuProTyrIleHisProThrAsn 60

QY 181  CCTTGGCCCGCATGTCACCGCTGCGCGCGGTGCGACCTGCTGTCGCCAACCCGCC 240
DB 61  ProPheAlaAlaAspValValSerGlnProGlyAlaGlyThrArgProArgGlnProPro 80

QY 241  CGACCACTCGGCTCGGCTTGGCGTGACAGGCGCCAGCGCGCGCGTTCCTCAGCTGT 300
DB 81  ArgProLeu**SerAlaTrpArgAspGlnSerGlnArgProSerAlaAlaProArg 100

QY 301  AGACCTACCAACAGCTGGCGCGCGCTAACCGCGTGTGCTCCGCGCCCATGACACCCG 360
DB 101  ArgSerAlaProAlaGlyAlaAlaProLeuThrAlaValSerProAlaProAspThrAla 120

QY 361  CGAGTGCCTGATGTGACTCCGCGCGCCCATCTTGC CGCGCGCAGTAACTATCAACA 420
DB 121  ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140

QY 421  TCTCCCTTACCTCTCCGTTGCCCGCCAGCTAACCTGGTCTCTTATGCGCGCTCTT 480
DB 141  SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160

QY 481  AGTCGCTTTTACCCCTTCAGACGCGCACCAATACCATATATGCGCCACGGAAGTCT 540
DB 161  AsnProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180

QY 541  AATTATGCCCAGTACCGGTTGCCGTGCCAATCCGTTACGCGCGCTGTGTCCCAAT 600
DB 181  AsnTyrAlaGlnTyrArgValValArgAlaThrIleArgTyrArgProLeuValProAsn 200

QY 601  GCTGTGCGCGGTTACGCCATCTCCATCTCATTTCTGCGCACAGCACACCCCGGCG 660
DB 201  AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220

QY 661  TCCGTTGATATGAATCAATAACCTCGACGGATGTTTCGTATTTTATGTCAGCGCGCAT 720
DB 221  SerValAspMetAsnSerIleThrSer**AspValArgIleLeuValGlnProGlyIle 240

QY 721  GCCTCTGAGCTTGATCCCAAGTAGCGCCTACACTATCGTAACCAAGCGTGGCGCTCC 780
DB 241  AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260

QY 781  GTCGAGACCTCTGGGTGCTGAGGAGGAGTACTCTGCTGCTTGTATGCTTTGTCATA 840
DB 261  ValGluThrThrGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280

QY 841  CATGCTCACTCGTAAATCTTATATAATACACCTATACCGGTGCGCTCGCGGTGTTG 900
DB 281  HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300

QY 901  GACTTGGCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACGGGGTTC 960
```

RESULT 15

US-10-319-745-167

; Sequence 167, Application US/10319745

```
DB 301  AspPheAlaLeuGluLeuGluPheArgAsnLeuLeuThrProGlyAsnThrAsnThrArgVal 320
QY 961  TCCCGTTATTCAGACACTGCTCGCCACCGCCTTCGTCGGGTGCGACCGGACTCGCGAG 1020
DB 321  SerArgTyrThrSerThrAlaArgHisArgLeuArgArgGlyAlaAspGlyThrAlaGlu 340
QY 1021  CTCACCAACACCGCTCTACCGCTTATGAAGGACCTCTATTATTTACTAGTACTAATGGT 1080
DB 341  LeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheAlaGlyThrAsnGly 360
QY 1081  GTCGGTGAGATCGCGCGCGGATAGCCTCACCGCTTCAACCTTGTCAACTGTGACACTGCTT 1140
DB 361  ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
QY 1141  GCGCGCTCGCCACACAAATTGATTCGTCGGTGGTGGCCAGCTGTCTACTCCCGTCCC 1200
DB 381  GlyGlyLeuProThrGlnLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
QY 1201  GTTGTCTCAGCCCAATGGCGAGCGGACTGTTAAGTTGTATACATCTGTAGAGATGCTCAG 1260
DB 401  ValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGln 420
QY 1261  CAGGATTAAGGGTATTGCAATCCGCGATGACATTCGACCTCGAGAAATCTCGTGTGGTATT 1320
DB 421  GlnAspLysGlyIleThrIleProHisAspIleAspLeuGlyAspSerArgValValIle 440
QY 1321  CAGGATTATGATAACCAATGAAACAAGATCGCGCGCGCTTCTCCAGCCCATCGCGC 1380
DB 441  GlnAspTyrAspAsnGln**GluGlnAspArgProThrProSerProAlaProSerArg 460
QY 1381  CCTTCTCTGCTCTCGAGCTAATGATGTGCTTTGGCTCTCTCTCACCGCTGCCAGTAT 1440
DB 461  ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyr 480
QY 1441  GACAGTCCACTTATGGCTCTTCGACTGGCCCGAGTTATGTTTCTGACTCTCTGACCTTGG 1500
DB 481  AspGlnThrThrTyrGlySerSerThrAsnProMetTyrValSerAspThrValThrLeu 500
QY 1501  GTTAATGTTGCCACCGCGCGAGCGCTTGCCTGCGCTCGATGTCGATGGACCGAGTCA 1560
DB 501  ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 520
QY 1561  CTTGACGGTTCGCGCCCTCTCCACCATCCAGCAGTACTCAAGACCTTCTTGTCTCGCG 1620
DB 521  LeuAspGlyArgProLeuThrThrIleGlnTyrSerLysThrPheTyrValLeuPro 540
QY 1621  CTCGCGGTAAGCTCTCTTCTGGGAGGAGCAGCACAATAAGCCGGGTACCTTATAAT 1680
DB 541  LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsn 560
QY 1681  TATAACACCACTGTAGCGACCAACTGTGTCGAGAGATGCGCGCGCGCGCGGTCGCT 1740
DB 561  TyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAla 580
QY 1741  ATTTCCACTTACCACTAGCCTGGGTGCTGCTCCCGTCTCCATTTCTGCGGTTCGCGTT 1800
DB 581  IleSerThrTyrThrThrSerLeuGlyAlaGlyProThrSerIleSerAlaValGlyVal 600
QY 1801  TTAGCCCCCACTCTGCGCTAGCTAGCTTGTGAGGATACCTTGAGTACCTGCGCGCGCC 1860
DB 601  LeuAlaProHisSerAlaLeuAlaValLeuGluAspThrIleAspTyrProAlaArgAla 620
QY 1861  CATACTTTTGTGATTTCTGCGCGAGTGC CGCGCTTGGCTTTCAGGGCTCGCGCTTTC 1920
DB 621  HisThrPheAspAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 640
QY 1921  CAGTCTACTGTCGCTAGCTTCAGCGCCTTAAGATGAAGGTGGGTAAACCTCGGAG 1977
DB 641  GlnSerThrIleAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu 659
```



```
; Publication No. US20030211467A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlaunder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232-US.P1
; CURRENT APPLICATION NUMBER: US/10/319,745
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: US/09/468,147A
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/173,141
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/061,199
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-15
; NUMBER OF SBO ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 167
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 84
; OTHER INFORMATION: Xaa = Unknown or Other at position 230
; OTHER INFORMATION: Xaa = Unknown or Other at position 447
; US-10-319-745-167

Alignment Scores:
Pred. No.: 5,63e-224 Length: 560
Score: 3151.00 Matches: 603
Percent Similarity: 95.45% Conservative: 26
Best Local Similarity: 91.50% Mismatches: 30
Query Match: 86.35% Indels: 0
DB: 12 Gaps: 0

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-319-745-167 (1-660)

Qy 1 ATCGCGCCCTGGCCCTATTGTTGCTGCTCTCTCATGTTTTTTCCTATGCTCGCGCGCCA 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1 MetArgProArgAlaValLeuLeuLeuPheValLeuLeuProMetLeuProAlaPro 20
Qy 61 CCGCGCGGTGACCGGTCTGGCGCGGTCTGGCGCGGTCTGGCGCGGTCTGGCGCGGTCTGG 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 21 ProAlaGlyGlnProSerGlyArgArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 40
Qy 121 TTCGGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCCCTATATTTCATCCACCAAC 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 41 PheTrpGlyAspArgValAspSerGlnProPheAlaLeuProTyrIleHisProThrAsn 60
Qy 181 CCCTCCCGCCCGGTGACCGGTCTGGCGCGGTCTGGCGCGGTCTGGCGCGGTCTGGCGCG 240
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 61 PropheAlaAlaAspValValSerGlnProGlyAlaGlyThrArgProArgGlnProPro 80
Qy 241 CGACCACTCGGCTCGGCTTGCGGTGACCGCGCGGTCTGGCGCGGTCTGGCGCGGTCTGG 300
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 81 ArgProLeu***SerAlaTrpArgAspGlnSerGlnArgProSerAlaAlaProArgArg 100
Qy 301 AGACCTACCACAGTGGGGCGCGGTCTGCAATCCCGGTCTGGCGCGGTCTGGCGCGGTCT 360
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 101 ArgSerAlaProAlaGlyAlaAlaProLeuThrAlaValSerProAlaProAspThrAla 120
Qy 361 CCAGTGCCTGATGTCGATCCCGCGCGGTCTGTCGCGCGGTCTGTCGCGCGGTCTGTCGCG 420
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
Qy 421 TCTCCCTTACTCTTCTCGGTGGCGACCGGCACTAACCTGGTCTTATGCGCGCGGTCTT 480
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
```

```
Qy 481 AGTCCGCTTTTACCCCTTCAGGCGCACCAATACCATATAATAGCGCACCAAGCTCTT 540
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 161 AsnProLeuLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
Qy 541 AATTATCCCAAGTACCGGGTTGCCGTCACCAATCCGTTACCCCGCGCTGGTCCCCAAT 600
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 181 AsnTyrAlaGlnTyrArgValValAlaThrIleArgTyrArgProLeuValProAsn 200
Qy 601 GCTGTCGGGGTTACGCGCATCTCATCTCTCTGTCGCGCACAGACACACACCCGACG 660
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
Qy 661 TCCGTTGATATGAATTCATAAATCCAGCATCTGCTATTTTATGTCACCCGCGCAT 720
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 221 SerValAspMetAsnSerIleThrSer***AspValargIleLeuValGlnProGlyIle 240
Qy 721 GCCTCTGAGCTTGATCCCAAGTGAGCGCTACACTATGTAACCAAGGTCGGCTCC 780
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSer 260
Qy 781 GTCGAGACCTCTGGGTGGCTGAGGAGGAGGCTACCTCTGCTGTTGTTATGCTTGCATA 840
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 261 ValGluThrThrGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
Qy 841 CATGGCTCACTCGTAAATTCCTATACATACACCTATACCGGTGCCCTCGGCTGTTG 900
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
Qy 901 GACTTTCCTTTCGAGTTGAGTTTCGCAACCTTACCCCGGTAAACCAATACCGGGTC 960
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrProGlyAsnThrAsnThrArgVal 320
Qy 961 TCCCGTTATTCACGACTGCTCGCCACCGCTTCGTCGCGGTGGCGGACGCGGACGCGAG 1020
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 321 SerArgTyrThrSerThrAlaArgHisArgLeuArgGlyAlaAspGlyThrAlaGlu 340
Qy 1021 CTCACACACGCGCTGCTACCGCTTTTATGAGGACCTCTATTTTACTAGTACTAATGCT 1080
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 341 LeuThrThrThrAlaAlaThrArgPheMetIleAspLeuHisPheAlaGlyThrAsnGly 360
Qy 1081 GTCGCTGAGATCGCGCGGATAGCCCTCACCTGTTCACCTGTGTCACCTGCTGCTGCT 1140
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 361 ValGlyGluValGlyArgGlyIleAlaLeuThrLeuPheAsnLeuAlaAspThrLeuLeu 380
Qy 1141 GCGCGCGCGCGCAGAAATGATTTCGTCGCGGTGGCGGCGGTCTCTACTCCGCGCC 1200
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 381 GlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgPro 400
Qy 1201 GTTCTCTCAGCCCAATGCGGAGCCGACTGTTAAAGTTGTATATCATCTGTAGAGAACTCAG 1260
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 401 ValValSerAlaAsnGlyGluProThrValIleLeuTyrThrSerValGluAsnAlaGln 420
Qy 1261 CAGGATTAAGGATTTGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGTGGTTATT 1320
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 421 GlnAspIysGlyIleThrIleProHisAspIleAspLeuGlyAspSerArgValIle 440
Qy 1321 CAGGATTTATTAACCAACATGACAGATCGCGGAGCTCTCTCCAGCCCGCATCCGCGC 1380
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 441 GlnAspTyrAspAsnGln***GluGlnAspArgProThrProSerProAlaProSerArg 460
Qy 1381 CTTTCTCTGCTCTTCGAGCTAATGATGCTGTTGGCTCTCTCTCACCGGTCCGCGAGTAT 1440
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 461 ProPheSerValLeuLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGlyTyr 480
Qy 1441 GACCAGTCCACTTATGGCTCTTCGACTGGCCCAAGTTTATGTTTCTGACTCTGTGACCTTG 1500
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 481 AspGlnThrThrTyrGlySerSerThrAsnProMetTyrValSerAspThrValThrLeu 500
Qy 1501 GTTAAATGTTGGACCGCGCGCGCGGTTCGCGGTGCGCTCGATTCGATTCGACCAAGTACA 1560
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 501 ValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThr 520
Qy 1561 CTTGACGCTCGCCCTCTCCACCATCCAGTACTCGAAGACCTTCTTTTGTCTCGCG 1620
```

```
Db      521 LeuAspGlyArgProLeuThrThrIleGlnTyrSerLysThrPheTyrValLeuPro 540
QY      1621 CTCGCGGTAAAGCTCTTTCTGGAGGAGGCAACAATAAGCGGGTACCCCTTATAAT 1680
Db      541 LeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAen 560
QY      1681 TATAACACCACTGCTAGCGACCAACTGCTTGTGAGATGCCCGGCGGACCGGGTCGCT 1740
Db      561 TyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAla 580
QY      1741 ATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGCTCCCATTTCTCGCGTTCGCGTT 1800
Db      581 IleSerThrTyrThrThrSerLeuGlyAlaGlyProThrSerIleSerAlaValGlyVal 600
QY      1801 TTAGCCCCCACTCTGCGCTAGCATTTGCTTGAGGATACCTTGAGTACCTGCGCCGCGCC 1860
Db      601 LeuAlaProHisSerAlaLeuAlaValLeuGluAspThrIleAspTyrProAlaArgAla 620
QY      1861 CATACTTTTGATGATTTCTGCCAGAGTGCCGCCCTTGGCCTTCAGGGCTGCGCTTC 1920
Db      621 HisThrPheAspAspPheCysProGluCysArgThrLeuGlyLeuGlnGlyCysAlaPhe 640
QY      1921 CAGTCTACTGTCGCTGAGCTTCAGCGCCTTAAGATGAAGTGGTAAACTCGGGAG 1977
Db      641 GlnSerThrIleAlaGluLeuGlnArgLeuLysMetLysValGlyLysThrArgGlu 659
```

Search completed: August 16, 2004, 13:33:17
Job time : 137.324 secs

QY 145 CAGCCCTTCGCAATCCCTATATTATCCACCA-----ACCCCTTC 186
Db 1157 gaIaProMetProThrProThrLeuValArgProLeuLeuLysLeuValHisSerProSe 1177
QY 187 GCGCCCGATGTACCGCTCGCGCGGGCTGGACCTCGTGTCCCAACCGCGCCGACCA 246
Db 1177 rProGluValSerAlaSerAlaProGlyAlaAlaProLeuThrIleSerSerProLeuHl 1197
QY 247 CTCGGCTCCGTTGGCGTACAGGCCCGCGCCCGCTGCTCAGCTCAGCTGATGACCT 306
Db 1197 s---ValProSerSerLeuProGlyProAlaSerSerProMetPro-----11 1212
QY 307 ACCACAGCTGGGCGCGCGCTAACCGCGCTG-----CTCCGGCC 348
Db 1212 eProAsnSerSerProLeuAlaSerProValSerThrValSerValProLeuSerSe 1232
QY 349 CATGACACCGCCGAGTGTCTGATGTGCTGCTCCCGCGCGCCATCTGGCGCGGAGTAT 408
Db 1232 rSerLeuProIleSerValProThrThrLeuProAlaProAlaSer-AlaProLeuThrI 1252
QY 409 AACCTATCAACATCTCCCTTACCTCTTCGGTGGCCACCGGCACCTAACCTGTTCTTTAT 468
Db 1252 leProIle-SerAlaProLeuThrValSerAlaSerGlyProAlaLeuLeuThrSerVal 1271
QY 469 GCGCCCTCTTAGTCCGCTTTTACCCCTTCAGGACGGCACCACATACCCATATAGGCC 528
Db 1272 ThrProProLeuAlaProValProAlaProAlaProGly-Pro-----1285
QY 529 ACGGAAGCTTCTAATTATGCCAGTACCGGGTTGCCCGCCACATCCGCTACCGCCCG 588
Db 1286 -----ProSerLeuAlaProSerGI 1292
QY 589 CTGTGTCCTCAATCTCTCGGGTTAG-----CATCTCCATCTCATTTCTGG 636
Db 1292 yAlaSerProSerAlaSerAlaLeuThrLeuGlyLeuAlaThrAlaProSerLeuSerSe 1312
QY 637 CCACAGACACCAACC-----ACCCGACGTCGCTGTGATGATGAATCAAT 680
Db 1312 rSerGlnThrProGlyHisProLeuLeuAlaProThrSerSerHisValProGlyLe 1332
QY 681 AACCTCGAGGATGTTTCGTATTTAGTCCAGCGCGGATAGCTCTGAGCTTGTGATCCC 740
Db 1332 uAsnSerThr-----ValAlaProAla---CysSerProValLeuValPr 1346
QY 741 AGTGAGCGCTAC-----A 755
Db 1346 oAla-SerAlaLeuAlaSerProPheProSerAlaProAsnProAlaProAlaGlnAlaS 1366
QY 756 CTATCGTAACCAAGGTGGCTCGCTCGAGACCTCTGGGGTGGCTGAGGAGGAGCTAC 815
Db 1366 eLeuLeuAlaProAlaSerSerAlaSerGlnAlaLeuAla-----ThrP 1381
QY 816 CTCGTGCTGTGTTATGCTTTGCATACATGCTCACTCGTAAATTCCTATACTAATACACC 875
Db 1381 roLeuAlaProMetAlaAlaProGlnThrAla-----IleLeuAlaP 1395
QY 876 CTATACCGTGCTCGCTGGGCTGTGGACTTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 935
Db 1395 roSerProAlaPro-----ProLeu-----AlaProLeuP 1405
QY 936 CCGCGGTAAACACCAATACGCGGTCTCCGTTATTCAGCACTGCTCGCACCGCTTCG 995
Db 1405 roValLeuAlaPro-----SerProGlyAlaAlaProValLeuAlaSerSerGlnT 1422
QY 996 TCGCGGTGGCGGAGCTCGCGAGTCTACCAACCGCGTG-----1036
Db 1422 hrProValProValMetAlaProSerSerThrProGlyThrSerLeuAlaSerAlaSerP 1442
QY 1037 -----CTACCGCTTTATGAGGACCTCTATTTTACTAGTACTAATGGTGT 1082
Db 1442 roValProAlaProThrProValLeuAlaProSerSerThrGlnThrMetLeuProAlaP 1462

QY 1083 CGGTGAGATCGCGCGGGGATAGCCCTCACCTGTGTTCAACCTT----- 1125
Db 1462 roValProSerProLeu-----ProSerProAlaSerThrGlnThrLeuAlaLeuAlaP 1480
QY 1126 -----CTGACACTCTGCTGGCGGCTCGGACAAATGATTGTTCTTCGCTGGCTGGTG 1177
Db 1480 roAlaLeuAlaProThrLeuGlySerSerProSerGlnThrLeuSerLeuGlyThrG 1500
QY 1178 GCGAGCTGTTCTACTCCCGTCCGCTGCTCAGCCAAATGGGAGCGCAGCTGTTAAATGTT 1237
Db 1500 lyAsnProGlnGlyProPheProThrGlnThrLeuSerLeuThrProAlaSerSerLeuV 1520
QY 1238 ATACATCTCTAGAGATGCTCAGCAGGATAAGGGTATTGCAATCCCGCATGACATTGACC 1297
Db 1520 alProThrProAlaGlnThr-----LeuSerLeuAlaProGlyProProL 1535
QY 1298 TCGGAGAATCTCGTGTGGTTATTTCAGGATTATGATAACCAACATGAACAGATCGGCCGA 1357
Db 1535 euGlyProThrGlnThrLeu-----SerLeuAlaProA 1546
QY 1358 CGCTTCTCAGCCCATCGCGCTTCTCTGCTCCTCGAGCTAATGATGCTTTGGC 1417
Db 1546 laProProLeuAlaProAlaSerProValGlyProAlaProAlaHisThr-----L 1563
QY 1418 TCTCTCTCAGCGTCCGAGTATGACAGTCCACTTATGCTCTCTCGACTGGCCAGTTT 1477
Db 1563 eutrLeuAlaProAla-SerSerSerAlaSerLeuLeuAlaProAlaSerValGlnThr 1582
QY 1478 ATGTTTCTGACTCTGACCTTTGTTAAATGTTGCGACCGCGCGGCTTGCCTGGT 1537
Db 1583 LeuThrLeuSer-----ProAlaProVal 1590
QY 1538 CGCTCGATTGGACCAAGTCACTTGACGGTGGCGCCCTCTCCACCATCCAGCAGTACT 1597
Db 1591 ProThrLeuGlyProAlaAlaAlaGlnThrLeuAlaLeuAlaPro-----AlaSerThr 1608
QY 1598 CGAAGACCTTCTTGTCTCGCTCGCGTCCGCGTAAAGCTCTTCTCGGAGCGGCGACAA 1657
Db 1609 GlnSerProAlaSerGlnAlaSerSerLeuValSerAlaSerGly-----AlaAla 1626
QY 1658 CTAAAGCGGGTACCTTATAATTATACCACTGCTAGCAGCAACCACTGTTGTCGAGA 1717
Db 1627 ProLeuProValThrMetValSerArgLeuPro-----ValSerLys 1640
QY 1718 ATGCGCGCGGACCGGCTGCTATTTCACCTACCACTAGCTAGCTGGTGGTGGTCCCG 1777
Db 1641 AspGluProAspThr----- 1645
QY 1778 TCTCCATTTCTCGGTTGCGGTTTAGCCCCCCTCTCGCTAGCATTTGCTTGAGGATA 1837
Db 1646 -----LeuThrLeu-ArgSerGlyProProSerProProSerThrAlaThrSerPh 1662
QY 1838 CTTTGAGTACCTCGCGCGCCCACTATTTTGATGATTTCTGCCAGAGTGGCGCCCC 1897
Db 1662 edGly-----ProArgPro-----ArgArgGlnProProPr 1673
QY 1898 TTGGCTTTCAGGCTGCGCTTTCAGTCTAC 1928
Db 1673 oProPro-----ArgSerProPheTyr 1680

RESULT 2

US-10-170-205E-18914
; Sequence 18914, Application US/10170205E
; GENERAL INFORMATION:

; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18914

```

; LENGTH: 1878
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-18914
Alignment Scores:
Pred. No.: 1.44e-05 Length: 1878
Score: 203.00 Matches: 163
Percent Similarity: 37.48% Conservative: 93
Best Local Similarity: 23.87% Mismatches: 269
Query Match: 5.56% Indels: 158
DB: 6 Gaps: 28

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-170-205E-18914 (1-1878)
Qy 13 CCTATTGTTGCTGCTCTCATGTTTGGCTATGTCGGCGCGCCACCGCCGGTCAG 72
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 224 ProValIleSerAlaAspSerProPheValProValIleProAla-AspGlyProVal-- 242
Qy 73 CCGTCTGGCCCGCTCGTGGCGGCGGAGCGGCTTCCGGCGTG---GTTTCTGGGGT 129
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 243 -----ValValAspSerSerThAspValIleValValProValPheProVa 258
Qy 130 GACCGGTTGATTCATCGACCTTCGCAATCCCTATATTCACCAACCAACCCCTTCGCC 189
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 258 lAsnGlyIleLeuVal-----ProAlaPheAlaAspGluThrValValPr 273
Qy 190 C-----CGATGTCACCGTGGCGGCGGCGGCTGGACCTCGTGTTCGC 231
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 273 oLeuAsnProAlaAspSerProValIleProValThrProAspAspLeuValValPr 293
Qy 232 CAACCCGCCGACCACTCGGCTCGGCTGGGTGACACGAGCCAGCGCCGCGCTGCC 291
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 293 oValThrProThrAspCysProValAlaProValThrSerAlaAspCysProValValPr 313
Qy 292 TCAC-----GTGTTAGACTACACAGCTGGGGCGCGCGCTAACCC 333
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 313 oPheThrProAlaAspAspProIleVal--ProValThrProAlaAspSerProIleVal 332
Qy 334 GCGTTCGCTCGGCGCCATGACACCCGCGAGTGCCTGATGTCAGTCCGCGGCGCCATC 393
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 333 SerValThrProAla--AspSerProValValPro----- 343
Qy 394 TTGCGCGGCGAGTAACTATCACTATCCCTTACCTCTTCGTTGGCCACCGGCACT 453
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 344 -----ValThrProAlaAspSerProValValProValThrProAlaAspGly 359
Qy 454 AACCTGGTCTTTATGCGCGCCCTCTTAGTCCGCTTTTACCCCTTCAGGACGGCACCAAT 513
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 360 ProValValProValAlaSerAlaAspSerProPheValProValThrProAlaAspAsp 379
Qy 514 ACCATATAATGGCCAGGAGCTCTTAATATGCCCAGTACCGGTTGCGCGTCCACACA 573
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 380 ProIleValSerValThrProAlaGluSerProAla--IleProValThrProAlaAsps 399
Qy 574 ATCGTTACCGCCGCTGTCGCCAATGCTGTCGGCGGTAGCCCATCTCCATCTCATTC 633
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 399 erProValValProValThrProAlaAspSerProValIle-SerValAlaSerAlaAsp 418
Qy 634 TGGCCACAGACACCACCCAGCCGCTCGGTTGATATGAATTAACCTTCGACGGAT 693
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 419 SerProValValProValThrProAlaAspThrValValValProValThrProAlaAsp 438
Qy 694 GTTCGTATTTAGTCCAGCCCGGCTAGCCTCTGAGCTTGTGATCCCAAGTAGCGCCTA 753
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 439 GlyProValIlePro-----ValThrProAlaGlu-SerPr 450
Qy 754 CACTATC---GTAACCAAGCTGGCGCTCCGTCGAGACCTCTGGGTGGCTCAGGAGGAG 810
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 450 oAlaIleProValThrProAlaAspSerProValValProVal----- 464
Qy 811 GCTAGCTCTGGTCTTGTATGCTTTGCATACATGGCTCACTCGTAAATTCCTATATAAT 870
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

465 -----ThrProValAlaAsps 469
Qy 871 ACACCTATACCGGTG---CCCTCGGCTGTTGGACTTTGCCCTTGAGCTTGAGTTTCGC 927
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 469 pProValValProValThrProThrAspPheProValValAlaValThrProAlaAspSe 489
Qy 928 AACCTTACCCCGGTAAACACCAATACCGGCTCTCCGTTATTATCCAGCACTGCTCGCCAC 987
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 489 rThrValAspProValThrProAlaAspGly--ProValValProVal----- 504
Qy 988 GCGCTTCGTCGCGTGGCGGAGCTGCGGAGCTACCCACAGCTGCTACCGCTTACCGCTTT 1047
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 505 -----ThrProAlaAspSerPro----- 510
Qy 1048 ATGAAGGACCTCTATTATTACTAGTACTAATGGTTCGGTGAGATCGGCGCGGATAGCC 1107
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 511 -----ValIleSerValThrProAla-AspAspProI 521
Qy 1108 CTCACCTGTTCAACCTTGTGACACTCTGCTTGGCGGCTG-----CCG 1152
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 521 leValProValThrPro-AlaAspSerIleuIleValSerValThrProAlaAspSerPro 540
Qy 1153 ACAGATTGATTCGTCGGCTGGCGAGCTGTTCTACTCCCGTCCGTTGCTCTCAGCC 1212
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 541 ValValSerValThrProAlaAspGlyArgValIle-----ProValAlaSerAla 557
Qy 1213 AATGGCGAG-----CCGACTGTTAAGTTGTATACATCTGTAGAGATGCTCAGCAGGAT 1266
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 558 AspSerProIleValProValThrProAlaAspThrValValValProValThrProAla 577
Qy 1267 AAGGCTATTGCAATCCCGCATGACATTGACCTCGGAGAACTCTCGTGTGTT---ATTGAG 1323
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 578 AspGlyLeuValIlePro-----ValThrAlaAlaGluCysProAlaIleProValThr 595
Qy 1324 GATTATGATAACCAACATGAACAGATCGCGGCGGCTTCTCCA----- 1368
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 596 ProValAspSerProValValProValThrSerAlaAspSerProLeuIleProValThr 615
Qy 1369 ---GCCCATCGCGCCCTTCTCTGCTCTGAGCTAATGATGCTTGTGGCTCTCTC 1425
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 616 ThrAlaAspSerArgAlaValProValThrProAlaGlyAspProIle---ValProVal 634
Qy 1426 ACCGCTGCGGAGTAGACCACTGCTTATGGCTCTTCGACTCGCCAGTCTTAT 1479
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 635 ThrProAlaAspThrProGlySerProValAlaSerAlaAspGlyProValProVal 654
Qy 1480 -----GTTTCTGACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGGCGGCTTGGC 1533
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 655 ThrProAlaAspGlyProValProValThrProAlaAspAspProValProVal 674
Qy 1534 CGGTGCTCGATTGACCAAGTCACTGACGCTCGCCCTCTCCA-----CCATC 1587
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 675 ThrSerAlaAsp-GlyProValValSerIleThrProAlaAspSerProValPheProVa 694
Qy 1588 CAGCAGTACTCGAGACCTCTTCTGCTGCGCTCCGCGGTAAAGTCTT-----CT 1638
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 694 lThrSerAlaAspGlyProValPheSerValThrSerAlaAspSerSerValProVa 714
Qy 1639 TTCTGGGAGGAGGAGCACAACATAAGCCG-----GGTACCTTTAATAT 1683
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 714 lThrSerAlaAspSerSerValPheProValThrThrAlaAspGlyProValProVa 734
Qy 1684 AACACCA-----CTGCTAGCAGCAACTGCTTGTGAGAAATGCCCGCG-- 1726
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 734 lThrProAlaAsnAspProValValLeuValThrSerAlaAspSerProValAlaProVa 754
Qy 1727 -----GCCACCGGCTCGCTATTTCACCTTACCACTAGCTGGTGGTGGTCCCGTC 1779
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 754 lSerSerAlaAspGlyProIleValProValThrProAlaAspSerLeuValSerVa 774
Qy 1780 TCCATTCTCGGGTGGCGGTTTATAGCCCCCACTCTGCGCTAGCATTCGTTGAGGATACC 1839
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Dy 774 lThrThrAlaAspAspProIle-----ValPr 783
```

```
QY 1840 TTGAGTACCTGCCCGCGCCCATCTTTGATGATTTCTGCCAGAGT 1888
Db 783 oValThrProAlaAspSerProValLeuSerValThrLeuAlaAspSer 799

RESULT 3
US-60-581-351-7551
; Sequence 7551, Application US/60581351
; GENERAL INFORMATION:
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated With Protein In Plants
; FILE REFERENCE: 38-21 (53372)B
; CURRENT APPLICATION NUMBER: US/60/581,351
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: US 60/479,962
; PRIOR FILING DATE: 2003-06-19
; NUMBER OF SEQ ID NOS: 13980
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7551
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-60-581-351-7551

Alignment Scores:
Pred. No.: 1-52e-05 Length: 763
Score: 201.00 Matches: 163
Percent Similarity: 32.03% Conservative: 58
Best Local Similarity: 23.62% Mismatches: 222
Query Match: 5.51% Indels: 248
DB: 7 Gaps: 37

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-60-581-351-7551 (1-763)
QY 7 CTCGGCCATTTTGTGTCCTCTCATGTTTTCCTATGTCGCCGGCCACGCC 66
Db 34 ProAlaProCysTrpMetLeuValSerAlaAlaProProCysProProAlaPro 53
QY 67 GGTGAGCCGTCTGGCGCGCGTGTGGCGCGGCGGCGGCGGTTCCGGTCTCG 126
Db 54 AlaProProLysProLysSerLysAla----- 62
QY 127 GGTGACCGGGTTGATTTCTCAGCCCTTC---GCAATCCCTATATTCATCAACCAACCC 183
Db 63 -----ProPheProProValProProAlaProProAlaArgGlu 75
QY 184 TTCGCCCCGATGTACCGCTGCGCGCGGGCTGACCTGCTGTCGCCAACCGGCCGA 243
Db 76 LeuAlaProProLeuProProAlaProProGluAlaProArgGluSerArgProAlaLeu 95
QY 244 CCACTCGGCTCG-----CTTGGCGTGACGAGCGCCGAGCGCCGCC 285
Db 96 Pro-ProCysProProProProValValleProAspProProGluProAlaAlaPro 115
QY 286 GTTGCTACGCTGTAGACTACCAAGCTGGGGCGCG-----CGCTAACCGCGTCCG 341
Db 115 oValProProAlaProAsnSerProProPheProProPheProAlaProLysPheVa 135
QY 342 TCGGCCCATGACACCCCGCGAGTGTGATGTGACTCCCGGGGGCCATCTTGGCGG 401
Db 135 lProAlaProProValProProAlaProAsn-SerProProPheProPro----- 151
QY 402 GCAGTATAACCTATCAATCTCCCTTACCTCTTCGTTGGCCACCGGCACTAACCTGT 461
Db 152 -----PheProProAlaAlaLeuAsnProProAlaProProAla----- 164
QY 462 TCTTTATGCGCCCTCTTAGTCCGCTTTTACCCCTTACGACGGCACAATACCATAT 521
Db 165 -----ProProLeuAla-----AsnSerProProLeuPro----- 174
QY 522 AATGGCCAGGAAGCTTCTAATTATGCCAGTACCGGGTTGCCCGTCCCAATCCGTTA 581
Db 175 -----ProAlaProProThrProAlaGlyThrProProAla 186

QY 582 CGCCCGCTGCTCCCAATGCTCTCGCGGTACGCCATCTCCATCTCATTTCTGGCCACA 641
Db 186 laAlaProTrpProProAlaProAlaAlaProLysSerLysProAla---SerProPro 205
QY 642 GACCACACACCCCGACGTCGCTTATATGAATTCATTAACCTCGACGGATTTCTGTAT 701
Db 205 xgProProAlaProProMetProAlaThrProMetGluPheProProLeuProProVal- 224
QY 702 TTTAGTCCAGCCCGCATAGCTCTGAGCTTGATGCCAAGTGAGCGGCTACACATCG 761
Db 225 -----ProProAspProLysSer----- 230
QY 762 TAACCAAGGCTGGCGCTCGTCCGAGACCTCTGGGTGGCTGAGGAGGAGCTACCTCTGG 821
Db 231 -----LysGluThrProProAlaProProAlaPro----- 240
QY 822 TCTTGTATGCTTTGCATACATGGCTCACTCGTAATTCCTATPACTAATACACCTATAC 881
Db 241 -----ProLysProProAlaProValP 248
QY 882 CGGTGCCCTCGGGCTGTTGGACTTTCCTGAGTTGAGTTTCGCAACCTTACCCCGG 941
Db 248 roilePro-----ProValProProL 255
QY 942 TAACACCAATACCGGGTCT-----CCCGTTATTCAGCACTGCTCGCCA 986
Db 255 euProProValProAsnLyslleProProAlaProProAlaProProValAlaAla 275
QY 987 CCGCTTCTGTCGGGTGCGGAGCGGACTCGCGAGCTCCGAGCTCACCA-----C 1028
Db 275 laValLeuValAlaProCysProProLeuProProLeuProAsnAsnHisProProAlaP 295
QY 1029 CAGGGCTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGTTCGGTGA 1088
Db 295 ro-ProAlaAlaProValProGlyValProLeu----- 305
QY 1089 GATCGCGCGGATGACCTCACCCTGTTCAACCTTGTGACACTCTGTTGGCGGCT 1148
Db 306 ---AlaProLeuProAsnSerHisProProAlaPro-----Pro 317
QY 1149 GCCGACAGATTTGATTTCTCGCTGGTGGTGGCAGCTGTTCTACTCCG-----TCCCGT 1202
Db 318 SerAlaProValProGlyVal-----ProLeuAlaProLeuProLysSerGlyArg 334
QY 1203 TGTCTCAGCAATGCGGACCGACTGTTAAGTTGTA-----TACTCTGTAGAGATCG 1256
Db 335 ProValSerValTrpLysGlySerPheThrThrLeuSerThrPheCysCysArgValCys 354
QY 1257 TCAGCAGGA-----TAAGGTATTGCAATCCGATGACATGACCTCGGAGATCTCG 1310
Db 355 SerGlyGluValLeuAlaGlyAlaLeuAsnProSer-----ArgProSerArg----- 370
QY 1311 TGTGTTATTACAGATTATGATTAACCAACATGACGCGGACGCTTCTTCCAGC 1370
Db 371 -----SerProLeu-----ThrThrThrProAla-LeuProAlaProL 384
QY 1371 CCCATCG---CGCCCTTCTCTGCTCGACCTAATGATGCTTGGCTCTCTCTCAC 1427
Db 384 eProProLeuProProLeuProProLysProLysAsnThrAlaVal-ProProLysPro 404
QY 1428 CGCTCCGAGTATGACCTGCTTATGCTCTTTCGACTCTTTCGAGTCTTATGTTTCTGA 1487
Db 404 roLeuProProValThrAla----- 410
QY 1488 CTCTGTACCTTGGTTAATGTTGACCGGCGGCGGCGGCTGCTCGATG 1547
Db 411 -----LeuAlaPro-----ProLeuPro----- 416
QY 1548 GACCAAGGTACACTTGGCGTCCCTCT-----CCACCAT 1586
Db 417 --ProLeuAlaProLeuProLysSerProGlyValProProAlaProProLysProPro 436
```

```
QY 1587 CCAGCAGTACTCGAAGACCTCTTTGTCCTCGCGCTCCGGGTAAGCTCTCTTTCTGCGA 1646
Db 436 lyIysProTprThrProLeuAla 445
QY 1647 GGCAGGCACAACTAAAGCCGGGTACCTTATAATATAACACCACTGCTAGCGACCAACT 1706
Db 446 --ProAlaProProGluProLysThrValProValLeuProPro-----GlyProSerC 463
QY 1707 GCTTGTGAGAAATCCGCGCGGACCGGTGCTATTTCACCTTACACCA----- 1756
Db 463 ys-----ProProSerGluLysProAsnProProAlaProProGluProProG 479
QY 1757 -----CTAGCCTGGGTGCTGCTCGCTCCA-----TTTC 1787
Db 479 luProLysSerSerProAlaLeuProProAlaProProAlaProSerMetProSerAlav 499
QY 1788 TGCAGGTTCGGCTTTAGCCCCCCTACTCTGCGCTAGCATGCTTGAGGATACCTTTGGACTA 1847
Db 499 alArgValProProSerProPro-----IlePro----- 508
QY 1848 CCCTGCGCGCGCCATACTTTTGATGATTCTGCCAGAGTCCGCGCCCTTGGCCTTCA 1907
Db 509 -----ProAlaPro-----ProAlaAlaProA 516
QY 1908 GGGCTGCGCTTCCAGTCTA 1927
Db 516 rgAlaSerMetProAlaLeu 522

RESULT 4
US-10-170-205E-19858
; Sequence 19858, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19858
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-19858

Alignment Scores:
Pred. No.: 2.25e-05 Length: 1006
Score: 199.00 Matches: 199
Percent Similarity: 30.35% Conservative: 55
Best Local Similarity: 23.78% Mismatches: 242
Query Match: 5.45% Indels: 341
DB: 6 Gaps: 44

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-170-205E-19858 (1-1006)
QY 34 ATGTTTTCCTATGCTCCGCGCGCCACCGCGGTGAGCGTCTGGCGCGGTGCTG-- 91
Db 137 MetPheHisThrValSerProGlyProProSerAlaArgProProCysArg-ValProPr 156
QY 92 -----GGCGGCGACGCGCGTTCGC----- 112
Db 156 oThrThrProLeuAsnGlyGlyProGlySerLeuProProGluProProSerValSerGl 176
QY 113 -----GGGTGGTTCCTGGGGTGACCGGGTG-----ATTCTCAGCCCT-- 151
Db 176 nAlaPheProThrLeuAlaGlyProGlyGlyLeuPheProProArgLeuAlaAspProVa 196
QY 152 -----TCGCAATCCCTATATTCATC-----CAACCAACCCCTTC 186
Db 196 lProSerGlyGlySerSerProArgPheLeuProArgGlyAsnAlaProSerProAl 216
QY 187 GCCCCCGATGTCACCG-----CT 204
```

```
Db 216 aProProProProAlaIleSerLeuAsnAlaProSerTyrAsnTrpGlyAlaAlaLe 236
QY 205 GCGCCCGGGGTG-----GACCTCGTGTTCGCCAACCCGCCCGCCACCACTCGGTCC 255
Db 236 uArgSerSerLeuValProSerAspLeu-----GlySerProPro-----AlaPr 251
QY 256 G--CTTGGGTGACACGAGCCCGCGCGCTTGCCTCAGCTCGTAGACCTA----- 307
Db 251 oHisAlaSerSerSerProProSerAspProProLeuPheHisCysSerAspAlaLeuTh 271
QY 308 ----CCACAGCTGGGGCG----- 322
Db 271 rProProProLeuProProSerAsnAsnLeuProAlaHisProGlyProAlaSerGlnPr 291
QY 323 -----CGCCGCTAACCGCGGTGCTCGCGCCCATGACACCCCGCCA 363
Db 291 oProValSerSerAlaThrMetHisLeuProLeuValLeuGlyProLeuGlyAlaPr 311
QY 364 GTGCTGATGTCGACTCCCGCGCGCATCTTGGCGCGGAGTATA----- 409
Db 311 oThrValGluGlyProGlyAlaProProPheLeuAlaSerSerLeuSerAlaAla 331
QY 410 -ACCTATCAACATCTCCCTTACCTCTTCG----- 439
Db 331 aLysAlaGlnHisProProLeuProProProSerThrLeuGlnGlyArgArgProArgAl 351
QY 440 -----TGCCACCGGCACCTAACCTGCTCTTTAT 468
Db 351 aGlnAlaProSerAlaSerHisSerSerSerLeuArgProSerGlnArgArgProArgAr 371
QY 469 GCGCCCTCTTAGTCGCGCTTTTACCCTTCAG----- 502
Db 371 gProProThrValPheArgLeuLeuGlyArgGlyProGlnThrProArgArgSerAr 391
QY 503 ----ACGCGCAATAACCATATAATGGCCACGGAAGCTTCTAATTATGCCCCAGTACCG 558
Db 391 gProArgAlaProAlaProValProGlnProPheSerLeu-----ProGluProSe 408
QY 559 GTTCCCGGTGCCACAATCCGTTACCGCGCTGTCGCCAATGCTGTCGGGGTTTACGCC 618
Db 408 rGlnProIleLeuProSerValLeuSer-----LeuLeuGlyLeuProThrPr 424
QY 619 ATCTCCATCTCATCT-----GCCACACAGACACC 648
Db 424 oGlyProSerHisSerAspGlySerPheAsnLeuLeuGlySerAspAlaHisLeuProPr 444
QY 649 ACCA-----CCCGACGTCGTTGATATGAATCA 678
Db 444 oProProThrLeuSerSerGlySerProProGlnProArgHisProIle----- 460
QY 679 ATACCTCGACGATGTTGCTATTATTAGTCACGCGCGCATAGCTCTGAGCTTGTGATC 738
Db 461 -----GlnProSerLeuPro----- 465
QY 739 CCAAGTGAGCGCTACACTATCGTAACCAAGCTCGCGCTCGCGACACCTCGGGGTG 798
Db 466 -GlyThrThrSerGlySerLeuSerSerValProGlyAlaProAlaProProAlaAla-- 484
QY 799 GCTGAGGAGGAGGCTACCTCTGGTCTGTTTGTGCTTTCATCATCATGCTCCTCGTAAAT 858
Db 485 -----SerLysAlaProValValPro-----Se 492
QY 859 TCCTATATAATACACCT----- 877
Db 492 rProValLeuGlnSerProSerGluGlyLeuGlyMetGlyAlaGlyProAlaCysProLe 512
QY 878 -----ATACCGGTGCGCTCG----- 892
Db 512 uProProLeuAlaGlyGlyGluAlaPheProPheProSerProGluGlnGlyLeuAlaLe 532
QY 893 -----GGTGTGGACTTTGGCCTTGGAGTTT 924
```



```
QY 526 GCCACGGAGCTTCTAATTATGATCCAGTACCGGGTTGCCGTGCCCAATCCGTTACCGC 585
Db 160 -----ProAspProSerPro-----ProSerAlaProValAl 171
QY 586 CCGCTGGTCCCAATGCTGCGGGTTACGCCATCTCCATCTCTGCGCACAGACC 645
Db 171 aProTipProProLeu-----ProProLeuProAsnAsnHisProPr 185
QY 646 ACCACCCCGCGAGCTCGTGTGATGATGAATTCATAACCTCGAGGATGTCGTATTTTA 705
Db 185 oAlaProProSerAlaProValProGlyVal----- 195
QY 706 GTCACGCGCGCATAGCTCTGAGCTGTGTATCCCAAGTGAGCGCTACACTATCGTAAC 765
Db 196 -----ProLeuAlaProLeuProfile----- 202
QY 766 CAAGCTGGCGCTCCGTCGAGACCTCTGGGTGGCTGAGGAGGCTACCTCTGGTCTT 825
Db 203 ---SerGlyArgProValArg-----AlaTipValGlySerLeuIleAlaLeuArgIl 219
QY 826 GTTATGC-----TTTGCATACATGGCTCACTCGTAAATTCCTAT 864
Db 219 eCysCysCysArgValCysSerGlyValLeuAlaGlyAlaLeuAsnProSerArgProSe 239
QY 865 ACTAATACACCTATACCG-----GTGCCCTCGGCTGTGGACTTTGGCCCTT 912
Db 239 rSerCysProProLysProProAlaProAlaValProAlaGlyAlaProValProProLe 259
QY 913 GAGCTTGAGTTTCGCAACTTACCCCGGTA-----ACACCAATACCGGGTCTCC 963
Db 259 u-----ProProLeuProProLeuProLeuSerThrProLeuProAlaPr 275
QY 964 CGTTATTCAGCACTGTCGCGACCGCTTCGTGCGGTGCGGAGGACTGCGGAGCTC 1023
Db 275 oProLeuProProLeuProAla-----LeuProThrSe 286
QY 1024 ACCA-----CCACGGCTGTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAAT 1077
Db 286 iProGlyAlaPro-ProAlaProProValProProAlaProAlaLys----- 301
QY 1078 GGTGTGCGTGAGATCGCGCGGGATAGCCCTCACCTGTGTTCAACCTTGTCACACTCTG 1137
Db 302 -----AspProProAlaProProAlaProProAlaProLeuSerArgProA 318
QY 1138 CTTGGCGGCTGCGGACAGAAATGATTTTCGCGCTGTGGTGGCCAGCTCTTCTACTCCCGT 1197
Db 318 laPheProProAlaPro-----ProAla-----Pro- 326
QY 1198 CCCGTGTCTCAGCAATGGCGAGCGGACTGTTAAGTTGTATACATCTGTAGAGATGCT 1257
Db 327 Pro---AlaSerLysProSerProThrProProAlaProProGluProAsnAsnVal 345
QY 1258 CAGCAGGATAAGGGTATT-----GCAATCCCGCATGACATTGACCTCGGAGAAATCTCGT 1311
Db 346 ProProSerProProLeuProProAlaProProProSerGlyLeu----- 361
QY 1312 GTGGTTATTCAGGATTATGATACCAACATGACAAAGATCGCGCGACGCCCTTCTCCAGCC 1371
Db 362 -----AspProProLeuProProAlaPro 369
QY 1372 CCATCGCGCCCTTTCTCTGTCTCTTC-----GA 1398
Db 370 ProAlaAlaPro-ArgLeuSerMetProAlaSerProProAlaProProPheProProTh 389
QY 1399 GCTAATGATGTGCTTGGCTCTCTCTCACCGCTG----- 1432
Db 389 rLeuIleMetLeuValProProLeuProProValProProAlaProAsnSerProProGl 409
QY 1433 ----CCGAGTATGACAGTCCACTTATGCTCTTCGACTGGCCCGAGTTTATGTTCTGAC 1489
Db 409 uProProSerProProAlaPro----- 416
```

```
QY 1489 TCTGTGACCTTGGTTAATGTTGTGACCGGGCGCGAGCCGTTGCCGTCGCTCGATGG 1548
Db 417 -----ProProLysMetProAsnProProGlyProProValProProAlaProAsnSe 434
QY 1549 ACCAAGTTCACACTTGACGGTCGCCCTCTCCACCATTC-----AGAGTAC 1596
Db 434 rPro-----ProPheProProAspProProAlaProProValProAlaSerVa 450
QY 1597 TCGAAGACCTTCTTTGTCTCCGCTCGCGGTAAAGTCTCTTCTTGGAGGAGGACACA 1656
Db 450 lAlaProProAlaProProThrProProSerAlaAsnSerProProPheProProAlaPr 470
QY 1657 ACTAAGCGCGGTACCCCTTATAATTAACACCACTGCTAGCAGCACTCTTGTCTGAG 1716
Db 470 oProAlaProPro-----ValAlaProLysAlaAlaAlaAsn----- 482
QY 1717 AATCCCGCGGACCGGGTCTGCTATTTCCACTTACCACTAGCTGGGTGCTGGTCCC 1776
Db 483 ---ProProGly-----ProProThrProAlaAla-----Pr 492
QY 1777 GTCTCCATTTCTGCGGTTCGGTGTAGCCGCCCACTCTGCGCTAGCATTCCTTGAGGAT 1836
Db 492 oAsnSerMetProAlaAlaProProAlaProPro----- 503
QY 1837 ACCTTGAGCTACCCCTGCGCGGCCATCTACTTTTGTATGATTTCTGCCAGAGTGCGCCCC 1896
Db 504 -----AlaProProValProValLeuAlaLeuProProAlaProProAlaProPr 520
QY 1897 CTTG 1900
Db 520 oLeu 521

RESULT 6
US-60-568-073-1197
; Sequence 1197, Application US/60568073
; GENERAL INFORMATION:
; APPLICANT: Huang, Fei
; APPLICANT: Han, Xia
; APPLICANT: Shaw, Peter
; APPLICANT: Clark, Edwin
; TITLE OF INVENTION: COLON CANCER BIOMARKERS
; FILE REFERENCE: 10205 PSP
; CURRENT APPLICATION NUMBER: US/60/568,073
; CURRENT FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 1305
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1197
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-568-073-1197

Alignment Scores:
Pred. No.: 0.000109 Length: 5179
Score: 192.00 Matches: 149
Percent Similarity: 33.69% Conservative: 74
Best Local Similarity: 22.51% Mismatches: 261
Query Match: 5.28% Indels: 178
DB: 7 Gaps: 33

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-60-568-073-1197 (1-5179)

QY 143 CTCAGCCCTTCGCAATCCCTATATTCATCAACCA-----ACCCCTTCGCCCCCGC--- 193
Db 1508 IleThrProProAlaSerThrThrLeuProProThrThrThrProSerProProThr 1527
QY 194 -----ATGTCACCGCTGCGCGGGGCTCGACCTCGTG 226
Db 1528 ThrThrThrThrThrProProProThrThrThrProSerProProThrThrThrProle 1547
QY 227 TTCGCCAACCGCGACCACTCGGCTCCGCTT---GGCGTGACGAGCCCGAGCCCGC 283
Db 1548 -----ThrProProThrSerThrThrLeuProProThrThrThrProSerProPro 1565
```


RESULT 10

US-10-170-205E-18875
; Sequence 18875, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18875
; LENGTH: 5262
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-18875

Alignment Scores:

Pred. No.:	0.000527	Length:	5262
Score:	182.00	Matches:	138
Percent Similarity:	31.12%	Conservative:	54
Best Local Similarity:	22.37%	Mismatches:	176
Query Match:	4.99%	Indels:	249
DB:	6	Gaps:	31

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-170-205E-18875 (1-5262)

QY	149	CTTTGGCAATCCCTATATTCATCCAAACCCCTTCGCCCGGATGTCACGGTGGCG	208
DB	498	ProGluGluSerProLeu-----SerProProGluSerSerProPheSer	513
QY	209	CGGGGCTGAGCTCGTGTGCGCAACCGCGACCACTCGCTCCGCTT--GGCGTG	265
DB	514	Pro----LeuGluGluSerProLeuSerProProGluGluSerProSerProAlaLeu	532
QY	266	ACGAGCCGACGCGCGCGTGGCTCAGCTGCTAGACCTACCAAGCTGGGGCGCGC	325
DB	533	GluThrProLeuSerProPro-----ProGluAlaSerProLeu	545
QY	326	CGCTAACCGGTGCTCGGCGCCATGACACCCCGCGAGTGGCTGATGTCGACTCCGCG	385
DB	546	SerProProPheGluGluSerProLeuSerProProProGluGluSerProThrSerPro	565
QY	386	GGCGCATCTTGGCGCGGAGTATAACATCAACATCTCCCTTACTCTTCGCGGCCA	445
DB	566	ProProGlu--AlaSerArgLeuSerProProProGluGluSerProMetSerProPro	584
QY	446	CGGCACTAACCTGGTCTTTATGCGCGCCCTCTTAGTCCGC---TTTTACCCCTTCAGG	502
DB	585	ProGluGluSerProMetSerProProProGluAlaSerArgLeuPheProPheGlu	604
QY	503	ACGGCACCAATACCCATATAATGGCCACGGAAGCTTCTAATTATGCCAGTACCGGTTG	562
DB	605	GluSerProLeuSer-----	609
QY	563	CCCGTGCCACATCCGTTACCGCGCTGGTCCCAATGCTGTGCGCGGTTACGCCATCT	622
DB	610	ProProProGluGluSerProLeuSerProProProGluAlaSerArgLeuSerProPro	629
QY	623	CCATCTCATTTGCGCACAGACACACACCCCGGACGTCCTGTTGATGATGAATCAATAA	682
DB	630	ProGluAspSerProMetSerProProProGluGluSerProMet-----	644
QY	683	CCTCGACGATGTTGTTATTTAGTCCAGCCGGCATAGCTCTGAGCTTGTGATCCCAA	742
DB	645	-----SerPro-----ProProGluVal---SerArg	652
QY	743	GTGAGCGCCTACACTATCGTAACCAAGGTGGCGCTCCGTCGAGA-----	787
DB	653	LeuSerProLeuProValValSerArgLeuSerProProProGluGluSerProLeuSer	672
QY	788	-----CCTCTGGGTGGCTGAGG-----	805

DB	673	ProProAlaLeuSerProLeuGluGluLeuGluTyProPheGlyAlaLysGlyAspSer	692
QY	806	---AGGAGGCTACTCTGCTGTTGATGCTTGCATACATCGCTCACTCGTAAATTCCT	862
DB	693	AspProGluSerProLeuAla-----	701
QY	863	ATACTAATAACACCTATA--CCGGTGCCCTCGGGCTGTGGACTTTGCCCTTGAGCTTG	919
DB	702	IleLeuGluThrProIleSerProProGluAlaAsnCysThrAspProGluProVal	721
QY	920	AGTTTGGCAACCTTACCCCGGTAAACCAATAACCGGGTCTCCGGTTATTCAGCA---	976
DB	722	ProProMetIleLeuProPro--SerPro-----GlySerProValGlyProAlaSer	738
QY	977	---CTGCTGCCACCGCTTCGTCGGGTGCGAGCGGACTGCCGAGCTCACCAACACGG	1033
DB	739	ProIleLeu-----MetGluProLeuProProGlu	748
QY	1034	CTGCTACCGCTTTATGAAGGACCTCTATTTTACTAGTACTAATGGTGTGGTGAGATCG	1093
DB	749	-CysSerProLeuGluHisSerLeuVal-----	758
QY	1094	GCCGGGATAGCCCTCACCTGTTCAACCTTGTGACACTCTGCTTGGCGGCTGCCGA	1153
DB	759	-ProGluAenSerPro-----	763
QY	1154	CAGAATTGATTTGTCGGCTGGTGGCCAGCTGTTCTACTCCGTCCTGTTCTCAGCCA	1213
DB	764	-----ProSerGlnCys-SerProP	770
QY	1214	ATGGCGAGCGGACTGTTAAGTTGTATACATCTGTAGAAATGCTCAGCAGGATAAGGTA	1273
DB	770	roAlaLeuPro-----L	774
QY	1274	TTGCAATCCCGCATGACATTTGACCTCGGAGAACTCGTGTGTTATTCAGGATTAATA	1333
DB	774	eusValProSerProLeuSerProIleGlyLysValValGlyValSerAspGluAlaG	794
QY	1334	ACCAATCAACCAAGATCGG-----	1353
DB	794	luLeuHisGluMetGluThrGluLysValSerGluProGluCysProAlaLeuGluProS	814
QY	1354	-----CCGAGCGCTTCTCCAG	1369
DB	814	eAlaThrSerProLeuProSerProMetGlyAspLeuSerCysProAlaProSerProA	834
QY	1370	CCCCATGCGCCCTTCTCTGCTTCGAGCTAATGATGTGCTTGGCTCTCTCTCACCG	1429
DB	834	laPro-----	835
QY	1430	CTGCGGAGTATGACCAATCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTCTGACT	1489
DB	836	--AlaLeuAspAspPheSerGlyLeuGlyGluAspThrAlaProLeu-----AspG	852
QY	1490	CTGTGACCTTGGTTAATGTTGCGACCGCGCGCGAGCCCTTCCCGGTCGCTCGATTGA	1549
DB	852	lyLeasp-----AlaProGlySerGln-ProGluProGlyGlnThrProGly	867
QY	1550	CCAAGGTACACTTGACGGTGCCTCCCTCTCCACCATCCAGCAGTACTCGAAGACCTTCT	1609
DB	867	-----	867
QY	1610	TTGTCTCGCGTCCGCTAGCTCTCTTTCTGGGAGCAGGACACAAAGCCGGT	1669
DB	868	-----SerLeuAlaSerGluLeuLysGlySerProValLeuLeuAspProGlu	883
QY	1670	ACCCTTATAATTATAACCACTG---CTAGCGACCAACTGCTGTGCGAGATGCGCGC	1726
DB	884	GluLeuAlaProValThrProMetGluValTyProGluCys-----LysGlnThrAla	901
QY	1727	GGCACCGGTCGCTATT-----CCACTTACCACTAGCTCGGTGGTGTG	1771

Db 902 GlyGlnGlySerProCysGluGlnGluProArgAlaProValAla----- 918
QY 1772 GTCCGCTCTCCATTCTCGCGTTCGCGTTTATAGCCCCCACTCTG 1816
Db 919 ---ProThr-----ProProThrLeu 924
RESULT 11
US-10-896-891-303
; Sequence 303, Application US/10896891
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/10/896,891
; CURRENT FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: 60/231,328
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 303
; LENGTH: 5262
; TYPE: PRT
; ORGANISM: Human
US-10-896-891-303
Alignment Scores:
Pred. No.: 0.000527 Length: 5262
Score: 182.00 Matches: 138
Percent Similarity: 31.12% Conservative: 54
Best Local Similarity: 22.37% Mismatches: 176
Query Match: 4.99% Indels: 249
DB: 6 Gaps: 31
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-896-891-303 (1-5262)
QY 149 CTTTCGCAATCCCTATATTATCAACCAACCCCTTCGCCCGGATGTCACCGCTGGG 208
Db 498 ProGluGluSerProLeu-----SerProProGluSerSerProPheSer 513
QY 209 CCGGGCTGACCTCGTTCGCGAACCCGCGGACCACTCGGCTCCGCTT---GGCGTG 265
Db 514 Pro---LeuGluGluSerProLeuSerProProGluGluSerProSerProAlaLeu 532
QY 266 ACCAGGCCACGCGCCCGCTGTGCTCAGCTCGTAGACACACAGCTGGGCGCGC 325
Db 533 GluThrProLeuSerProPro-----ProGluAlaSerProLeu 545
QY 326 CGCTAACCGCGCTCGCTCCGCGCCATGACACCCCGCCAGTGCCTGATGCTGACTCCCGC 385
Db 546 SerProProPheGluGluSerProLeuSerProProProGluGluLeuProThrSerPro 565
QY 386 CGCGCATCTTCGGCGGAGATACCTATCAACATCTCCCGTTACCTCTCCGTTGGCCA 445
Db 566 ProProGlu---AlaSerArgLeuSerProProProGluGluSerProMetSerPro 584
QY 446 CCGCACTAACTGGTTCTTTATGCGCGCTCTTAGTCCG---TTTACCCCTTCAGG 502
Db 585 ProGluGluSerProMetSerProProProGluAlaSerArgLeuPheProPheGlu 604
QY 503 ACGGACCAATACCATATAATATGGCCACGGAAGCTTCTAATTATGCCAGTACCGGGTTG 562
Db 605 GluSerProLeuSer----- 609
QY 563 CCGTGCCCAATCCGTTACCGCGCTGTCGCCCAATGCTGTGCGCGGTTACGCCATCT 622
Db 610 ProProProGluGluSerProLeuSerProProProGluAlaSerArgLeuSerProPro 629
QY 623 CCATCTCATCTGGCCACAGACCCACCCACCGTCCGTTGATATGAATCAATAA 682
Db 630 ProGluAspSerProMetSerProProProGluGluSerProMet----- 644
QY 683 CCTCGACGGATGTTCTGTATTTTAGTCCAGCCGCGCATAGCCTCTGAGCTTGTGATCCCA 742

Db 645 -----SerPro-----ProProGluVal---SerArg 652
QY 743 GTGAGCGCTACACTATCTGTAACCAAGCTGGCGTCCGTCGAGA----- 787
Db 653 LeuSerProLeuProValValSerArgLeuSerProProProGluGluSerProLeuSer 672
QY 788 -----CCTCTGGGTGGCTGAGG----- 805
Db 673 ProProAlaLeuSerProLeuGlyGluLeuGluTyProPheGlyAlaLysGlyAspSer 692
QY 806 ---AGGAGGCTACCTTGTCTTGTATGCTTTGCATACATGGCTCACTCGTAAATTCCT 862
Db 693 AspProGluSerProLeuAla-----AlaPro 701
QY 863 ATACTAATACACCCCTATA---CCGCTGCCCTCGGCTGTTGGACTTTCGCCCTTGAGCTTG 919
Db 702 IleLeuGluThrProIleSerProProGluAlaAsnCysThrAspProGluProVal 721
QY 920 AGTTTCGCAACCTTACCCCGGTAAACACAATACGCGGTCTCCGTTATTCCAGCA--- 976
Db 722 ProProMetIleLeuProPro---SerPro-----GlySerProValGlyProAlaSer 738
QY 977 ---CTGCTCGCCACCGCTTCGTCGCGTGCAGCGGACTCCGAGCTCCAGAGCTCACACACGG 1033
Db 739 ProIleLeu-----MetGluProLeuProProGln 748
QY 1034 CTGCTACCGCTTTATGAAGACCTCTATTTTACTAGTACTAATGGTTCGCTGAGATCG 1093
Db 749 -CysSerProLeuGluGlnHisSerLeuVal----- 758
QY 1094 GCCGCGGATAGCCCTCACCCCTGTTCAACCTTGTGACACTCTGCTTGGCGCCTGCCGA 1153
Db 759 -ProGlnAsnSerPro----- 763
QY 1154 CAGAATTGATTTCGTCGCTGTCGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTG 1213
Db 764 -----ProSerGlnCys-SerProp 770
QY 1214 ATGGGAGCCGACTGTTAAGTTGTATACATCTGTAGAGAAAGTCTCAGCAGGATAGGTA 1273
Db 770 roAlaLeuPro-----L 774
QY 1274 TTGCAATCCGCTACATGACCTCGGAGAATCTCGTGTGTTATTTCAGGATTATGATA 1333
Db 774 euSerValProSerProLeuSerProIleGlyLysValValGlyValSerAspGluAlaG 794
QY 1334 ACCAACATGACAAAGATCGG----- 1353
Db 794 luLeuHisGluMetGluThrGluLysValSerGluProGluCysProAlaLeuGluProS 814
QY 1354 -----CCGACGCTTCTCCAG 1369
Db 814 erAlaThrSerProLeuProSerProMetGlyAspLeuSerCysProAlaProSerProA 834
QY 1370 CCCCATCGCGCTTCTCTCTCTCGAGCTAATGATGCTTGTGGCTCTCTCTCACCG 1429
Db 834 laPro----- 835
QY 1430 CTGCGGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCAGTTTATGTTTGTGACT 1489
Db 836 --AlaLeuAspAspPheSerGlyLeuGlyGluAspThrAlaProLeu-----AspG 852
QY 1490 CTGTGACCTTGTAAATGTTGCGACCGCGCGAGCGGTTCGCCGTCGCTCGATTGGA 1549
Db 852 lyIleAsp-----AlaProGlySerGln-ProGluProGlyGlnThrProGly 867
QY 1550 CCAAGGTACACTTGACGGTTCGCCCTCTCCACCATCCAGCTACTCTGAAGACCTTCT 1609
Db 867 ----- 867
QY 1610 TTGCTCTCGCGGTAAAGCTCTCTTTTCTGGGAGCGGACCACTAAAGCCGGT 1669

```
Db 868 -----SerLeuAlaSerGluLeuLysGlySerProValLeuLeuAspProGlu 883
QY 1670 ACCTTTATATATACACACTG-----CTAGCGACCAACTGTTGTGCGAGATGCCGCG 1726
Db 884 GluLeuAlaProValThrProMetGluValThrProGluCys-----LysGlnThrAla 901
QY 1727 GGCACCGCGTGCCTAATT-----CCACTTTACACCACTAGCTGGGTGGTGG 1771
Db 902 GlyArgGlySerProCysGluGluGlnGluProArgAlaProValAla----- 918
QY 1772 GTCCCGTCTCCATTCTCGCGTTCGCGTTTGTAGCCCCCACTGTG 1816
Db 919 ---ProThr-----ProProThrLeu 924

RESULT 12
US-10-851-438-79
; Sequence 79, Application US/10851438
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Hertman, John L.
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-750
; CURRENT APPLICATION NUMBER: US/10/851,438
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/825,751
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/194,314
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 60/225,693
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Gallus gallus
US-10-851-438-79

Alignment Scores:
Pred. No.: 0.000428 Length: 1151
Score: 180.50 Matches: 163
Percent Similarity: 31.23% Conservative: 65
Best Local Similarity: 22.33% Mismatches: 258
Query Match: 4.95% Indels: 244
DB: Gaps: 33

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-851-438-79 (1-1151)
QY 80 GCCCGCGTGGCGCGCGCGGTTTCGGCGGTGGTTCGGGGTGACCGGGTTG 139
Db 392 AlaAlaThrValProThrAlaGlyAlaValProLysAlaSerThrGlyThr----- 409
QY 140 ATTCTCAGCCCTCGCAATCCCTATTATCATCCACCAACC----- 181
Db 410 -----ProAlaAlaPro-----GlnGlnProValProLysAlaAlaProVal 424
QY 182 -----CCATTGCGCC----- 190
Db 425 ThrProProSerProGlnGlnAlaValProArgAlaAlaThrAlaAlaAlaProVal 444
QY 191 ---CCGATGTCACCGTGGCGCGGGCTGGACCTCGTGTTCGCCAACCCGCCGACCAC 247
Db 445 ThrProGlnGlnProValThrLysAlaAlaThrThrAsnAlaThrProProGln 464
QY 248 TCGGCTCCGCTGGCGTGACCGCCAGCGCCCGCGTGGCTCAGCTGACCTA 307
Db 465 ProfileProLysAlaAlaThrThrThrAlaThrProValThrProGlnGlnProfile 484
```

```
QY 308 CCACAGCTGGCGCGCGCGCTAAACCGCGGTGCTCCGGCCCATGACACCCCGCAGTGC 367
Db 485 ProLysAlaGlyThrAspAlaAlaProProProAlaValProLysAlaProSerAspGly 504
QY 368 CTGATGTCACCTCCCGCGCGCATCTTCGGCGCGGAGTATACCTATCAACATCTCCCCC 427
Db 505 ArgAlaAlaThrProGlyValProAsnAlaAla-----ThrAspProGlnLysPro 521
QY 428 TTACTCTTTCGTTGG-----CCACGGCACTAACTGTTCTTTATGGCGCCCTCTTA 481
Db 522 ProProThrProGlnSerValProSerAlaValThrGluProLysProGlnPro----- 539
QY 482 GTCCGCTTTTACCCCTTCAGGACGCCACCA-----ATA 514
Db 540 -----ArgAlaAlaProProProSerAsnGluAlaThrProAlaVal 553
QY 515 CCCATATAATGGCCACGGAAGCTTCTAATTATATGCCAGTACCGGGTTGCCGTG----- 568
Db 554 ProSerProSerProAsnLeuLysSerProLeuProThrIleProLysProValProLeu 573
QY 569 -----CCACAATCCGTTACCGCGCGTGTCCCAATGCTGTCGGGGTTACG 616
Db 574 MetAlaLeuThrProGlnProValThrAlaGlnMetValThrGlnLeuAlaAlaThrLys 593
QY 617 CCATCTCCA-----TCTCATTTCTGGCCACAGACACCCACCCCGACGT 661
Db 594 ProSerProIleValProLysAlaSerProLysAlaLeuMetThrProProProPro 613
QY 662 CCGTTGATATGAATTCAATACTCGA-----CGGATGTCGTTATTTAG 706
Db 614 ProGlyLeu-----ProArgAlaLeuAlaAlaLysLeuGlyLeuPro 629
QY 707 TCAGCCCG----- 715
Db 630 SerSerProValAlaSerAlaMetHisAlaLysValThrProArgProLeuProAlaSer 649
QY 716 -----GCATAGCTCTGAGCTTGATCCCAAGTAGCGCTACACTATC 760
Db 650 ProValProMetAlaAlaSerProAlaSerLeuGlyProAspAlaAlaArgValAlaLeu 669
QY 761 GTAACCAAGCTGGCGCTCCGTCGAGACCTCTGGGTGGCTGAGGAGGCTACCTCTG 820
Db 670 AlaThrAsnAlaAlaSerPro-----Gly--AlaLysProGluAlaAlaGlyG 685
QY 821 GTCTTGTATGCTTTCATACATGCTCCTCCTGTA-----AATTCCTACTA 868
Db 685 ly-----AsnGlyThrLeuMetAlaProMetGlyAlaAlaAsnThrG 699
QY 869 ATACACCCCTATACCGGTGCCCTCGGCTGTTGGACTTTCCTGCTGAGCTTGAGTTTCGCA 928
Db 699 lnMetAlaProIleGlyAlaAlaGlyAlaAlaGlnThrAla----- 712
QY 929 ACCTTACCCCGGTAAACCAATACGGGTCTCCGTTTCCTGATTCACGACTGTCGCCACC 988
Db 713 -----ProMetGlyAlaAlaHisThrHisValSer----- 722
QY 989 GCCTTCGTGGCGTGGCGGAGCTGCCGAGCTCACCACTACCGGCTGCTACCGCTTTA 1048
Db 723 -----ProMetGlyAlaGlyGlyAlaThrGlnMetSerProThrGlyAla----- 737
QY 1049 TGAAGGACCTCTATTTTACTACTAATGTTGTTGCTGAGATCGGCGCGGATAGCCC 1108
Db 738 -----AlaAsnThrHisMetSerProIleGlyAlaGlyGlyAla 751
QY 1109 TCACCCCTG-----TTCAACCTTGTGACACTCTGTTGGCGGCTGCCGA 1153
Db 751 hrGlnMetSerProMetGlyAlaAlaAsnThrGlnMetSerProMetGlyAlaThrThr 771
QY 1154 CAGAAATTGATTTTCGTCGGCTGGTGGCCAGCTGTTCTACTCCCGTCCCGTTGCTCAGCCA 1213
Db 771 hrGlnMetSerProMetGlyAlaAlaAlaThrThrGlnProSerProMetGlyAlaAla 791
```



```
QY 1214 ATGGCGAGCGCGTGTAAAGTTGTATATCATCTGTAGAGAATGCTCAGCAG-----GATA 1267
      ::: ||||| ||||| |||||
Db 791 laThrGlnValThrAla-----ThrSerAlaGlyAsnThrMetGlnValSerProM 808
      ||||| ||||| ||||| |||||
QY 1268 AGGGTATTGCAATCCGCGATGACATTGACCTCGGAGAAATCTCGTGGTTATTACGGATT 1327
      ||||| ||||| ||||| |||||
Db 808 etGlyAlaAlaThrProGln-----
      ||||| ||||| ||||| |||||
QY 1328 ATGATAACCAACATGAACAAGATCGCGCAGCGCTTCT-----CCAG 1369
      ||||| ||||| ||||| |||||
Db 816 -----ThrProSerValGlyAlaAlaThrThrPro 826
      ||||| ||||| ||||| |||||
QY 1370 CCCATCGCGCCCTTCTCTGTCCTCGAGCTAATGATGTGCTTGGCTCTCTC----- 1423
      ||||| ||||| ||||| |||||
Db 826 InProSer-ProMetGlyAlaAlaThrThrLeuMetSerProMetGlyAlaAlaThrThr 845
      ||||| ||||| ||||| |||||
QY 1424 -----TCACCGCTGCCGAGTATGACAGT-----CCACTTATGGCTCTTGA 1465
      ||||| ||||| ||||| |||||
Db 846 ProGlnProSerProMetGlyAlaValThrThrGlnProProProMetAlaAlaThrAsn 865
      ||||| ||||| ||||| |||||
QY 1466 CTGGCCCGAGTTTATGTTTCTGACTCTGTGACCTTGTGTTAATGTTGCGACCGCGCGCAGG 1525
      ||||| ||||| ||||| |||||
Db 866 ThrThrGln-----
      ||||| ||||| ||||| |||||
QY 1526 CGTTGCGCGGTCGCTCGAATGGACCAAGTCAACATTTGACGCTGCGCCCTCTCCACCA 1585
      ||||| ||||| ||||| |||||
Db 869 ProProMetAlaAlaSerThrProGlnSer-----ThrProMetGlyAla 884
      ||||| ||||| ||||| |||||
QY 1586 TCCAGCAGTACTCGAAGACCTCTTTGTCCTCGCGCTCGCGGTAAAGTCTCTTCTGGG 1645
      ||||| ||||| ||||| |||||
Db 885 AlaThrThrThrGlnSerProProMetGlyAlaThrThrThrGlnSerProProMetGly 904
      ||||| ||||| ||||| |||||
QY 1646 AGGCGGCACAACTAAAGCCGGTACCCCTT-----ATA 1678
      ||||| ||||| ||||| |||||
Db 905 AlaSerThrProGlnAlaProProThrValAlaGlySerProThrProProProProle 924
      ||||| ||||| ||||| |||||
QY 1679 ATTATACACCACTGCTAGCGACCAAC-----TGCCTGTGAGAAATGCGCGCGGCACC 1732
      ||||| ||||| ||||| |||||
Db 925 ProProSerProThrAlaGlnThrSerProGlnProMetSerLysSerProProProAsp 944
      ||||| ||||| ||||| |||||
QY 1733 GGGTCGCTATTCCA-----CTTACACCACTAGCTGGGTGCTGTGTCGG 1777
      ||||| ||||| ||||| |||||
Db 945 ProProLysAlaProSerAlaAlaGlnThrSerProAlaAlaHisValAlaAsnAla 964
      ||||| ||||| ||||| |||||
QY 1778 TCTCCATTCTGCGGTGCGGTTTGTAGCCCTTGTAGCTGATTTCTGCCAGAGTGCCTCCGCC 1837
      ||||| ||||| ||||| |||||
Db 965 SerProGlyValThr-----
      ||||| ||||| ||||| |||||
QY 1838 COTTGGACTACCTGCGCGGCCCATACTTTTGTATGATTTCTGCCAGAGTGCCTCCGCC 1897
      ||||| ||||| ||||| |||||
Db 970 -----AlaValSerProAlaPro-----IleGlyValThrGluAlaSerPro 983
      ||||| ||||| ||||| |||||
QY 1898 TTGGCTTCAGGGCTGCGCTTTCC 1921
      ||||| ||||| ||||| |||||
Db 984 SerAlaAspGlyAlaArgLeuSer 991
      ||||| ||||| ||||| |||||
RESULT 13
US-10-425-115-313976
; Sequence 313976, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 313976
; LENGTH: 708
; TYPE: PRT
```

```
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(708)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_49406C.1.pep
US-10-425-115-313976

Alignment Scores:
Pred. No.: 0.000402 Length: 708
Score: 180.00 Matches: 155
Percent Similarity: 32.72% Conservative: 59
Best Local Similarity: 23.70% Mismatches: 231
Query Match: 4.93% Indels: 210
DB: 6 Gaps: 28

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-425-115-313976 (1-708)

QY 41 TGCCTATGTCGCCGCGCACCGCCGCTCAGCCGCTCTGGCGCGCTCGTGGCGCGCGCA 100
      ||||| ||||| ||||| |||||
Db 5 CysArgAlaLeuGlnPheHisProAlaAla-----Ser 16
      ||||| ||||| ||||| |||||
QY 101 GCGCGGTTCCGCGGCTGTTTCTGGGTGACCGGGTTGATTCTCAGCCCTTCGCAATCC 160
      ||||| ||||| ||||| |||||
Db 17 ThrProProSerLeuValLys-----LeuSerProProGlnAla 30
      ||||| ||||| ||||| |||||
QY 161 CCTATTATCATCAACCAACCCCTTCGCCCGCGATGTACCGCTGCGCGCGGGGTGGAC 220
      ||||| ||||| ||||| |||||
Db 31 ProVal-----GlySerProProProValLysThrThrSerProProAlaPro 47
      ||||| ||||| ||||| |||||
QY 221 CTCGGTTCGCCCAACCGCCGACCACTCGGCTCGCTGGCTGACCGAGGCCAGCGCC 280
      ||||| ||||| ||||| |||||
Db 48 IleAlaSerProSerProPro-----ProProAlaProValValSerProPro--P 64
      ||||| ||||| ||||| |||||
QY 281 CGCGCGTTGCTCAGTCGTAGACTACACAGCTGGGGCGCGCGCTAAACCGCGGTG 340
      ||||| ||||| ||||| |||||
Db 64 toProValLysSerProProProAlaProValGlySerProProProProGluLys 84
      ||||| ||||| ||||| |||||
QY 341 CTCGGGCCCATGACACCGCGCGAGTCTGTATGTCGACTCCCGCGCGGCATCTTTCGCC 400
      ||||| ||||| ||||| |||||
Db 84 erPro-----ProProProAlaPro-ValAlaSerPro----- 94
      ||||| ||||| ||||| |||||
QY 401 GSCAGTATAACCTATCAACATCTCCCTTACCTCTTCGTTGGCCACCGCSACTAACCTGG 460
      ||||| ||||| ||||| |||||
Db 95 -----ProProProValLysSerProProProProThrLeu 106
      ||||| ||||| ||||| |||||
QY 461 TTCCTTTATGCCGCCCTCTTAGTCGCTTTTACCCCTTCAGGACGCGCAACCAATACC 520
      ||||| ||||| ||||| |||||
Db 107 ValAlaSerProProProValLysSerProPro-----ProProAlaProVal 123
      ||||| ||||| ||||| |||||
QY 521 TAATGGCCACGGAAGCTTCTAATTATGCCAGTACCGGGTTCGCCGTG----- 568
      ||||| ||||| ||||| |||||
Db 124 AlaSerProProProValLysSerProProProProProProProProProProPro 143
      ||||| ||||| ||||| |||||
QY 569 ---CCACAATCGTTACCGCGCTGGTCCCAATGCTGTGCGCGGTTACGCCATCTCCA 625
      ||||| ||||| ||||| |||||
Db 144 ProProAlaProValAlaSerSerProProProMetLysSerProProProProPro 163
      ||||| ||||| ||||| |||||
QY 626 TCTCATTTCTGCCACAGACCA-----CCACACCCCGCAGCTCCGTTGATGAATT 676
      ||||| ||||| ||||| |||||
Db 164 ValSerSerProProProGluLysSerProProProProProProProProProPro 183
      ||||| ||||| ||||| |||||
QY 677 CAATAACTTCGACGATGTTGTTATTTAGTCCAGCCCGCATACCTCTGAGCTTGCA 736
      ||||| ||||| ||||| |||||
Db 184 ProProProGluGluTy-----ProThrProProThrSerVal--- 196
      ||||| ||||| ||||| |||||
QY 737 TCCCAAGTGAGCGCCTACACTATCGTAACCAAGCTGGCGCTCCGCTCGAGACCTCTGGGG 796
      ||||| ||||| ||||| |||||
Db 197 -----LysSerSerProProPro----- 202
      ||||| ||||| ||||| |||||
QY 797 TGGCTGAGGAGGAGGCTACCTCTGCTTGTATTGCTTTGCATACATGGCTCACTCGTAA 856
      ||||| ||||| ||||| |||||
```

```
Db 203 -----GlulysSerLeuPro----- 207
Qy 857 ATTCCTATAATA-----CACCCCTATACCGTGCCTCGGCTGTGG 901
Db 208 ProProThrLeuIleProSerProProGluGluLysProThrProSerThrPro 227
Qy 902 ACITTCCTGAGTGTAGTTTCGCAACCTTACCCCGGTAAACACCAATACGCGGTCT 961
Db 228 SerLysProProSerSerProGluLysProSerProProLysGluProValSerPro 247
Qy 962 CCGGTTATTCCAGCACTGCTGCACCGCTTCGTCGGTGGCGGACGGAGTCCG--- 1018
Db 248 ProGluThrPro-LysSerSerProProAla--ProValSerSerProProThr 266
Qy 1019 -----AGTCACCAACCGGCTGCTACCGCTTTATGAAGGACCTCTATTTACTAGTA 1072
Db 267 ProValSerSerProProAlaLeuAlaProVal----- 277
Qy 1073 CTAATGGTGTGGTGAGATCGCGCGGGATAGCCCTCACCTGTTCAACCTTGTCTGACA 1132
Db 278 -----SerSerProProSer----- 282
Qy 1133 CTCTGCTTGGCGGCTGCCGACAGAAATTGTTGTCGGCTGTGGCCAGCTTTCTACT 1192
Db 283 -----VallysSerSerProPro 288
Qy 1193 CCCGTCCTGTTCTCAGCCAAAT-GGCGAGCCGACTGTTAAGTTGTATACATCTGTAGAG 1251
Db 289 ProAlaProLeuSerSerProProProAlaProGluValLys-----SerSerProPro 306
Qy 1252 AATGCTCAGCAGGATAAGGTTATGCAATCCCGCATGACATTGACCTCGGAGAAATCTCGT 1311
Db 307 ProValGluValSerSerProProProAlaPro-----LysSerSer 320
Qy 1312 GTGGTTATTGAGTATTGATACCAACATGACAGATCGCGGACGCTTCTTCAGCC 1371
Db 321 ProProLeuAlaProValSerSerProProGluValGluLysThrSerProProAla 340
Qy 1372 CCA-----TCGCGCCCTTCTCTGTCCTCGAGCT 1401
Db 341 ProLeuSerSerProProLeuAlaProLysSerSerProProHisValValSerSer 360
Qy 1402 AATGATGTGCTTGGCTCTCTCAGCGTGCAGT---ATGACCAGTCCACTTATGGC 1458
Db 361 ProProProValValLysSer-SerProProProAlaProValSerSerProProLeuTh 380
Qy 1459 TCTTGACTGCGCCAGTTTANGTTCTGACTCTGTGACCTTGTGTTAATGTGCGACCGC 1518
Db 380 rProLysProAlaSerProProAlaHisValSerSerProGluValValLysProSe 400
Qy 1519 GCGCAGGCGTTGCCCGTCTCGATTGGACCAAGTCACTTGAGGTCGCCCTC 1578
Db 400 rThrProProThrProThrThrValLysSerPro-----ProSe 413
Qy 1579 TCACCATCAGCAGTACTCGAAGACCTTTTGTCTCGCTCGCTCGCGTAAAGTCTCT 1638
Db 413 rGluProLysSerSer----- 418
Qy 1639 TTCTGGGAGGCGCACTAAAGCGGGTACCTTATAATTATACACACTGTAGC 1698
Db 419 -----ProProProThrProProValSerLeuProProIleVa 431
Qy 1699 GACCAACTGCTTGTGAGATGCGCGCGGACCGGCTCGCTATTTCCACTTACACCACT 1758
Db 431 Llys-----SerSerProProProAlaMetValSerSerProProMetThrProLy 448
Qy 1759 AGCCTGG-----GTGTGTTCCCGTCTCCATTTCGCGGTGCGGTTTATAGCCCC 1809
Db 448 sSerSerProProProValValSerSerPro-----ProPr 461
Qy 1810 CACTCTGGGTAGATTGCTTGAGATACCTTGAGACTACCTGCGCGCCCATACTTTT 1869
Db 461 oThrValLys-----SerSerProProProAlaPro----- 471
```

```
Qy 1870 GATGATTTCTGCCAGAGTGGCGGCCCC 1897
Db 472 ----ValSerSerProAlaThrPro 479
```

RESULT 14

```
US-10-425-115-260103
; Sequence 260103, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 260103
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(317)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_168832C.1.pep
US-10-425-115-260103
```

Alignment Scores:

Pred. No.:	0.000372	Length:	317
Score:	179.00	Matches:	85
Percent Similarity:	29.51%	Conservative:	18
Best Local Similarity:	24.36%	Mismatches:	105
Query Match:	4.91%	Indels:	141
DB:	6	Gaps:	18

US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-425-115-260103 (1-317)

```
Qy 43 CTATATGTCGCGCGCACGCGCGGTCTGCGCGCGTGTGGCGCGGCGGAGC 102
Db 2 ProMetProProProProProProPro-----ProAsn 15
Qy 103 GCGGTTCCGGCGGTGTTCTGGGGTGACCGGGTTGATTCAGCCCTTCGCAATCCC 162
Db 16 -ProLeuPro-----LeuProAsnProLysProAsnPr 26
Qy 163 TATATTCAATCAACCAACCCCTTTCGCGCGGTGTACCGGTGCGCGGGGTGGACCT 222
Db 26 oLeuProAsnProProProProProProProProProProProProArgProPr 46
Qy 223 CGGTTCGCAACCGCGCGACCACTCGGCTCGTTCGGTGGCGTACACAGGCCGACG 277
Db 46 oProLeuProAsnProAlaProMetProCysProLysProProProProProPr 66
Qy 278 -----GCCCGCGGTGCTCAGTCGTAGACTACCA---CAGTGGGCGCGCGC 327
Db 66 oProArgProProProLeuProAsnProAlaProMetProCysProLeuProPro----- 84
Qy 328 CTAAACGCGGTCTCGCGCCATGACACCCCGCCAGTGTGCTGATGTGCTGACTCCGCGC 387
Db 85 ----ProArgProProProProAlaPro***ProLeuProProProArgProPr 103
Qy 388 GCGA-----TCTTGGCGCGGAGTATACCTATCAATCATCTCCC 426
Db 103 oProLeuProAsnProProProMetSerCysPro-----IlePr 116
Qy 427 CTTACCTCTTCGTCGCGCACCGCACTAACTCGTCTTTTATGCGCGCCCTCTTAGTCCG 486
Db 116 oProProLeuProProProProProMetProMetProProMetProProPro----- 132
```

```
QY 487 CTTTACCCTTCAGGACGGACCAATACCCATATAATGGCCACGGAAGCTTCTAATTAT 546
Db 133 -----SerProProLeuProTyPro----- 140
QY 547 GCCAGTACCGGGTTGCCGTGCCAATCCGTTACCGCCCGCTGGTCC-----CCAAT 600
Db 141 -ProProArgProSerProPhePro**ProProLeuMetProTyrSerIleProProPr 160
QY 601 GGTGTGGGGTTACGCATCTCCATCTCATCTCTGGCCACAGACCAACACCCCGGAG 660
Db 160 oLeuPro**ProProProMetPro-----TyrProMetProProProSerProPr 177
QY 661 TCCGTTGATATGAATTCATAACCTCGACGGATGTTGCTATTGTTAGTCCAGCCCGGCATA 720
Db 177 oProPhe-----Pro**Pr 182
QY 721 GCCTCTGAGCTTGTATCCCAAGTGAGCGCTPACACTATCGTAAACAGGTGGGCTCC 780
Db 182 oPro----- 183
QY 781 GTCGAGACTCTGGGGTGG---CTGAGGAGGAGGCTACTCTGGTCTTGTATGCTTTGC 837
Db 184 -----ProMetSerTyrProIleProProProLysPro----- 194
QY 838 ATACATGGCTCACTCGTAAATTCCTATATAACACCTATACCGGTGCCCTCGGGCTG 897
Db 195 -----ProProLeuProTyrProProProMetPro----- 204
QY 898 TTGGACTTTGGCCTTGAGCTTGAGTTTCGCAACCTTACCCCGGTAAACACCAATACGGG 957
Db 205 -TyrProIlePro-----ProProArgProPro----- 213
QY 958 GTCCTCCGTTATCCAGCATCTGCTGCCACCCGCTTCGTCGGGGTGGACGGGACTGCC 1017
Db 214 -----ProLeuPr 216
QY 1018 GAGCTCACCAACCGCTGCTACCC 1042
Db 216 o***ProProArgProTyrPro 224
RESULT 15
US-10-170-205E-16484
; Sequence 16484, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16484
; LENGTH: 1453
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-16484
Alignment Scores:
Pred. No.: 0.000734 Length: 1453
Score: 177.50 Matches: 185
Percent Similarity: 30.85% Conservative: 68
Best local Similarity: 22.56% Mismatches: 248
Query Match: 4.86% Indels: 319
DB: 6 Gaps: 44
US-09-851-410A-6_COPY_5147_7129 (1-1983) x US-10-170-205E-16484 (1-1453)
QY 37 TTTTGGCTATGCTGCCGCGC----- 57
Db 435 PheLeuProMetAspProAlaThrPheArgArgLysArgProGluSerValGlyGlyLeu 454
QY 58 CCACCGCCCGGT-----CAGCCGCTCTGGCCGCGCTGTCGTGGCGGCGC 99
```

```
Db 455 GluProProGlyProSerValIleAlaAlaProProSerGly----- 468
QY 100 AGCGCGGTTC----- 114
Db 469 ---GlyGlyAsnIleLeuGlnThrLeuValLeuProProAsnLysGluGluGlnGly 487
QY 115 GGTGGTTTCTGGGGTACCGGGTTGATTCTCAGCCCTTC----- 153
Db 488 Gly-----GlyAlaArgValProSerAlaProAlaProSerLeuAlaTyrGlyAla 504
QY 154 ---GCAATCCCTCATATTATCATCCA----- 174
Db 505 ProAlaAlaProLeuSerArgProAlaAlaThrMetValThrAsnValValArgProVal 524
QY 175 -----ACCAACCCCTTC----- 186
Db 525 SerSerThrProValProIleAlaSerLysProPheProThrSerGlyArgAlaGluAla 544
QY 187 GCCCGCGATGTACCGCTGCGGCC-----GGGCTGGACCTCGTCTTCGCCAA 234
Db 545 SerProAsnAspThrAlaGlyAlaArgThrGluMetGlyThrGlySerArgVal---Pro 563
QY 235 CCGCGCCGACCACTCGGCTCGGCTGGGTGACACGAGCCCGCCGCGCTTGCCTCA 294
Db 564 GlyGlySerProLeuGlyValSerLeuValTyrSerAspLysLysSerAlaAlaThr 583
QY 295 CCGCTGTAGACT-----ACCACAGTGGGGCGCGCGCGTCA 330
Db 584 SerProAlaProHisLeuValAlaGlyProLeuLeuGlyThrValGlyLysAlaProAla 603
QY 331 ACCGCGGTC-----GCTCGGCGCATGACACCCCG----- 360
Db 604 ThrValThrAsnLeuLeuValGlyThrProGlyTyrGlyAlaProAlaProProAlaVal 623
QY 360 ----- 360
Db 624 GlnPheIleAlaGlnGlyAlaProGlyGlyGlyThrThrAlaGlySerGlyAlaGlyAla 643
QY 361 -----CCAGTGCCT-----GATGTCGACTCC 381
Db 644 GlySerGlyProAsnGlyProValProLeuGlyIleLeuGlnProGlyAlaLeuGlyLys 663
QY 382 CGCGGGCCNCTTCGCGCGGAGTAACTATCAATCAATCAATCAATCAATCAATCAATCCG 441
Db 664 AlaGlyGlyIleThrGlnValGlnPyrIleLeuProThrLeuProGlnGlnLeuGlnVal 683
QY 442 -----GCCACCGGCACATACTACCTGGTCTTTATGCGCGCCCTCTTAGT 483
Db 684 AlaProAlaProAlaProAlaProGlyThrLysAla-----AlaAlaProSerGly 700
QY 484 CCG-----CTTTTACCCCTTCAGGACGGACGACCAATACCCAT 519
Db 701 ProAlaProThrThrSerIleArgPheThrLeuProProGlyThrSerThrAsnGlyLys 720
QY 520 ATAATGGCCACGGAAGCTTCTAATTATGCCAGTACCGGTTGCCGTCGCCAATCCGT 579
Db 721 ValLeu-----AlaAlaThr-AlaProThrProGlyIleProIleLeuGlnSerVa 737
QY 580 TACCGCGCGCTGCTCCCAATGCTGTCGGCGGTACGGCATCTCCATCTCATCTTCGCCCA 639
Db 737 lProSerAlaProProLysAlaGlnSerValSerProValGlnAla----- 753
QY 640 CAGACCAACACA-----CCCGACGCTCGGTTGATFANGAATCA 678
Db 754 ----ProProProGlyGlySerAlaGlnLeuLeuProGlyLysValLeuVal----- 769
QY 679 ATAACCTCGACGGATGTCGTATTTAGTCCAGCCCGGCATAGACCTCTGATGTC 738
Db 770 -----ProLeu----- 771
QY 739 CCAAGTGAGCGCTACACTATGCTAAACGAGCTGGCGCTCGTCGAGACCTCTCGGGTG 798
```

```
Db 772 -----AlaAlaProSerMetSerValArgGlyGlyValaGlyGlnPro----- 786
QY 799 GCTGAGGAGAGCTACCTCTGCTCTGTGTATCTTTGTCATACATGGCTCACTCGTAAAT 858
Db 787 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db -----LeuProLeuValSerProProPheSer-----Va 796
QY 859 TCCT-----ATACTAATACACCTATACCGGT 885
Db 796 lProValGlnAsnGlyAlaGlnProProSerLysIlelleGlnLeuThrProValProVa 816
QY 886 G-----CCCTCGGCTGTGGACTTTGCCCTTGAGCTTGAGTTTCGCAACCTTACCCCC 939
Db 816 lSerThrProSerGly-----LeuValProProLeuSerProAlaThrLeu--Pro 832
QY 940 GGTAAACACC-----AATACGGGGTCTCCCGTTAT 969
Db 833 GlyProThrSerGlnProGlnLysValLeuLeuProSerSerThrArgIleThrTyVal 852
QY 970 TCCAGCACTGCTCGCCACCGCTTCGTGCGGT-----GCGGACGGGACTGCCGAG 1020
Db 853 GlnSerAlaGlyHisAlaLeuProLeuGlyThrSerProAlaSerSerGlnAlaGly 872
QY 1021 CTCACACACAGGCTGTACCGCTTTATGAAGCACTCTATTACTAGTACTAATGGT 1080
Db 873 ThrValThrSerTyGlyProThrSerSerValAlaLeuGlyPheThrSer----- 889
QY 1081 GTCGGTGAGATCGCGCGGATAGCCCTCACCTGTTCAAC----- 1122
Db 890 LeuGlyProSerGlyProAlaPheValGlnProLeuLeuSerAlaGlyGlnAlaProLeu 909
QY 1123 CTTGCTGACACTCTGCTGGCGGCCTGCCGACA-----GAATTGATT 1164
Db 910 LeuAlaProGlyGlnValGlyValSerProValProSerProGlnLeuProAlaCys 929
QY 1165 TCCTCGCTGGTGGCGCAGCTG-----TTCTAC----- 1191
Db 930 AlaAlaProGlyGlyProValIleThrAlaPheTy-SerGlySerProAlaProThrSer 949
QY 1192 TCCCGTCCGCTTCTCTCAGCAATGGCGAGCGACTGTTAAGTTGTATACATCTGTAGAG 1251
Db 950 SerAlaProLeuAlaGlnProSerGlnAlaProProSerLeuValTyThrValAlaThr 969
QY 1252 AATGCTCAGCAGGATAGGGTATTGCAATCCCGCATGACATTGACCTCGGAGAACTCGT 1311
Db 970 SerThrThrProProAlaAlaThrIleLeuPro----- 980
QY 1312 GTGGTTATTGAGATTATGATACCAACATGACAGATCGGCGGAGCGCT----- 1362
Db 981 -----LysGlyProProAlaProAlaThrAla 989
QY 1363 TCTCCAGCCCATCGCGCCCTTTCTCTGCTCTCGAGCTAATGATGTGCTTTGGCTCTCT 1422
Db 990 ThrProAlaProThrSerProPheProSerAlaThrAlaGlySerMetThrTy--Ser 1008
QY 1423 CTCACGCTGCCAGATGACACAGTCCACTTATGGCTTCCTTCGACTGGCCCGCTTATGTT 1482
Db 1009 LeuValAlaProLysAlaGlnArgPro-----SerProLysAlaProGlnLysVal 1025
QY 1483 TCTGACTCTGTGACCTTGGTTATGTTGGACCGCGCGGCGGCGCTTGCCCGGTGCTC 1542
Db 1026 LysAlaAlaIleAlaSerIleProValGlySerPheGluAlaGlyAlaSer-GlyArgPr 1045
QY 1543 GATTGGACCAAGGTCACACTTGACGGTGGCCCGCTCTCCACCATCCAGCAGTACTCGAAG 1602
Db 1045 o----GlyProAlaProArgGlnProLeuGluProGlyPro-----ValArgGl 1060
QY 1603 ACTTCTTTGCTCGCGGTCCGCGGTAACTCTCTTTCTGGGAGGAGGAGGACAACTAAA 1662
Db 1060 uProThr-----AlaProGluSerGluLeuGluGlyGlnProThrProAl 1076
QY 1663 GCGGGTACCCCTT-----ATATTATACACC 1689
Db 1076 aProProProLeuProGluThrTrpThrProThrAlaArgSerSerProProLeuProPr 1096
```

```
QY 1690 ACTGCTAGCACCACCACTGCTTGTGAGAAATGCGCGGCGCACCGGGTCTGCTATTTCCTACT 1749
Db 1096 oProAlaGluGluArgThrSerAlaLysGlyProGluThrMetAlaSerLysPheProSe 1116
QY 1750 TACACCACTAGCCTGGGTGCTGCCGTCTCCATTTCTGCGGTTCGCGTTTGTAGCCCCC 1809
Db 1116 rSerSerSerAspTip---ArgValProGlyGlnGlyLeuGluAsnArgGlyGluProPr 1135
QY 1810 CACTCTGCGCTAGCATTTGCTTGAGGATACCTTGAGACTACCTGCGCGCGGCC 1861
Db 1135 oThr-----ProProSerProAlaProAlaPro 1144
```

Search completed: August 16, 2004, 13:27:00
Job time : 103.28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: August 16, 2004, 13:35:43 ; Search time 89.6821 Seconds
(without alignments)
12476.164 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096
Perfect score: 659
Sequence: 1 ATCGCCCTAGGCTCTTTT.....GTAAACTCGGAGTTGTAG 1980

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues
Word size: 1
Total number of hits satisfying chosen parameters: 2987630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09851410/runat_16082004_125716_9617/app_query.fasta_1.4238
-DB=A Geneseq 29Jan04 -QPMF=fastan -SUFFIX=oli.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09851410@cgn_1_141/runat_16082004_125716_9617 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A.Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	659	100.0	659	2 AAR38787	AAR38787 HEV ORF2
2	659	100.0	659	2 AAR35827	AAR35827 Hepatitis
3	572	86.8	659	2 AAR39308	AAR39308 Mexico st
4	558	84.7	659	2 AAR93387	AAR93387 Human HEV
5	512	77.7	660	2 AAR96090	AAR96090 Hepatitis
6	401	60.8	549	2 AAR96092	AAR96092 Hepatitis
7	392	59.5	540	2 AAR96102	AAR96102 Hepatitis
8	377	57.2	525	2 AAR96104	AAR96104 Hepatitis
9	327	49.6	327	2 AAR96094	AAR96094 Hepatitis
10	327	49.6	327	2 AAR35821	AAR35821 Hepatitis

11	309	46.9	659	5 ABJ04808	ABJ04808 Hepatitis
12	149	22.6	549	2 AAR76367	AAR76367 Hepatitis
13	149	22.6	660	2 AAR70323	AAR70323 Hepatitis
14	149	22.6	660	2 AAR91814	AAR91814 Hepatitis
15	149	22.6	660	2 AAR76369	AAR76369 Hepatitis
16	149	22.6	660	2 AAR93386	AAR93386 Human HEV
17	149	22.6	660	2 AAR93388	AAR93388 Human HEV
18	149	22.6	660	2 AAR93390	AAR93390 Human HEV
19	149	22.6	660	5 ABB81670	ABB81670 Hepatitis
20	149	22.6	660	6 ADA50060	ADA50060 Hepatitis
21	149	22.6	660	7 ADB97807	ADB97807 HEV ORF2
22	146	22.2	660	5 ABJ04820	ABJ04820 Hepatitis
23	146	22.2	660	5 ABJ04811	ABJ04811 Hepatitis
24	146	22.2	660	5 ABJ04807	ABJ04807 Hepatitis
25	146	22.2	660	5 ABJ04809	ABJ04809 Hepatitis
26	146	22.2	660	5 ABJ04819	ABJ04819 Hepatitis
27	146	22.2	660	5 ABJ04805	ABJ04805 Hepatitis
28	124	18.8	547	2 AAR26189	AAR26189 Epidemic
29	124	18.8	660	2 AAR93394	AAR93394 Human HEV
30	124	18.8	674	7 ADE06706	ADE06706 Hepatitis
31	124	18.8	674	7 ADE06704	ADE06704 Hepatitis
32	121	18.4	660	5 ABJ04822	ABJ04822 Hepatitis
33	121	18.4	660	5 ABJ04821	ABJ04821 Hepatitis
34	121	18.4	660	5 ABJ04813	ABJ04813 Hepatitis
35	121	18.4	660	5 ABJ04815	ABJ04815 Hepatitis
36	120	18.2	525	2 AAR96103	AAR96103 Hepatitis
37	120	18.2	540	2 AAR96101	AAR96101 Hepatitis
38	120	18.2	549	2 AAR96091	AAR96091 Hepatitis
39	120	18.2	660	2 AAR38785	AAR38785 HEV ORF2
40	120	18.2	660	2 AAR93306	AAR93306 Burma str
41	120	18.2	660	2 AAR96089	AAR96089 Hepatitis
42	120	18.2	660	2 AAR35826	AAR35826 Hepatitis
43	120	18.2	660	2 AAR80197	AAR80197 Protein e
44	120	18.2	660	2 AAR71210	AAR71210 Protein e
45	120	18.2	660	2 AAR93389	AAR93389 Human HEV

ALIGNMENTS

RESULT 1
AAR38787
ID AAR38787 standard; protein; 659 AA.
AC AAR38787;
XX
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 11-JAN-1994 (first entry)
XX
DE HEV ORF2 protein.
XX
XX Enterically transmitted non-A non-B hepatitis; ET-NANB; vaccine.
OS Hepatitis E virus; Mexico strain.
XX
XX Key Location/Qualifiers
FT Peptide 225..659
FT Peptide 333..659
FT Peptide 612..659
FT Peptide /label= 406.3-2
XX
XX WO9314116-A1.
XX
XX 22-JUL-1993.
XX
XX 15-JAN-1993; 93WO-US000459.
XX
XX 17-JAN-1992; 92US-00822335.
PR 01-MAY-1992; 92US-00876941.
XX
XX (GENE-) GENELABS TECHNOLOGIES INC.

PA	(USSH) US SEC DEPT HEALTH.
XX	
PI	Reyes GR, Bradley DW, Tam AW, Carl M;
XX	
DR	WPI; 1993-243144/30.
DR	N-PSDB; AAQ47130.
XX	
PT	New immunogenic hepatitis E virus (HEV) peptide(s) - are from the ORF1,
XX	ORF2 and ORF3 regions of HEV, useful as a vaccine against HEV infection.
XX	Disclosure; Fig 8; 48pp; English.
XX	
CC	Immunogenic hepatitis E virus (HEV) peptides are selected from the ORF1,
CC	ORF2 and ORF3 regions of HEV. The peptides can be used in vaccines to
CC	prevent infection by HEV. The antibodies can neutralise and block HEV
CC	infection and can be used to prevent or treat HEV infection. The peptides
CC	and antibodies can also be used as diagnostic reagents. (Updated on 25-
CC	MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS
CC	field)
XX	
SQ	Sequence 659 AA;
	Alignment Scores:
Pred. No.:	0 Length: 659
Score:	659.00 Matches: 659
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	2 Gaps: 0
US	-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR38787 (1-659)
QY	1 ATGCGCCTTAGGCCTTTTGGCTGTTCCTCTGTTCTTGCTATGTTGCCGCGCCA 60
Db	
1	MetArgProArgProLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY	61 CGACCGCGTCAGCGCTCTGGCCGCGCTCGTGGCGCGCACGCGCGGTACC GCGCGTGGT 120
Db	
21	ProThrGlyGlnProSerGlyArgArgGlyArgSerGlyGlyThrGlyGlyGly 40
QY	121 TTCTGGGCTGACGGGTGATTCTCAGCGCTTCGCATCCCTATATTCATCCACCAAC 180
Db	
41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrlIeHisProThrAsn 60
QY	181 CCCTTTGCCCCAGAGCTTGCCGCTCGCTCGGCTCGGACTCTGGACCTCGCCTTCGCCAACACAGCC 240
Db	
61	ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY	241 CCGCCACTTGGCTCCACTTGGCGAGATCAGGCCAGCCCTCCGCTGCTCCCGTCGC 300
Db	
81	ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
QY	301 CCACTCTGCCACAGCGGGCTGCGCGCTGAGCGGTGGCGCTCGCCCATGACACCTCA 360
Db	
101	ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHISAspThrSer 120
QY	361 CCGCHCCCGACGTTGATTCTCGCGGTGCCAATTCTACGCGCCAGTAGTATAATTTGTCTACT 420
Db	
121	ProValProAspValAspSerArgglyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY	421 TCACCCCTGACATCTCTGTGCGCTCTGGCACTAAATTAGTCTGTATGAGCGCCCCCTT 480
Db	
141	SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY	481 AATCCGCCTCTCGCGTCGAGACCGGTACTAATTAATCTCACATTATGGCCACAGAGGCTCC 540
Db	
161	AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY	541 AATTATGACAGTACCGGGTTCGCGCGCTACTATCGTTACCGGCCCTTAGTGCCTAAT 600
Db	
181	AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValProAsn 200
QY	601 GCAGTTGGAGGCTATGCTATATCCATTCTTCTCTGGGCTCAAAACAACACCAACCCCTACA 660


```
QY 901 GACTTTGCTTACAGCTTTCGAAATCTCACCACCTGTAAACCAATACACGCTGTG 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTACTCCAGCACTGCTCTACTCCGCGCGGAGGCGCGAGGAGCTGGAGCTG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCACAATCGACGCCACCAAGTTTCATGAAAGATCTCCACTTTTACCGGCTTAAATGCGGTA 1080
Db 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAAGTGGCGCGCGGAGTAGCTTAACATTACTTAACCTTGCTCACACGCTCTCCGCG 1140
Db 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GSGCTCCGACAGAAATTAATTCGTCGCTGGCGGGCAACTGTTTTATTCCCGCCCGGTT 1200
Db 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GTCTCAGCCAATGGCGAGCAACCGTGAAGCTCTATACATCAGTGGAGAAATGCTCAGCAG 1260
Db 401 ValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAGGGTGTGCTATCCCCACGATATCGATCTGGTGATTCGGCTGTGTCATTCAG 1320
Db 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGln 440
QY 1321 GATTATGACAACAGCAGCAGGATCGGCCACCCGCTCGGCTCGGCATCTCGGCT 1380
Db 441 AspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTTCTCCGAGCAAAATGATGTACTTTCGCTGCTCCTCACTCAGCCGAGTATGAC 1440
Db 461 PheSerValLeuArgAlaAsnAspValLeuTyrLeuSerLeuThrAlaAlaGluTyrAsp 480
QY 1441 CAGTCCACTACGGGTGCTCAATCGCCCGGTTTATATCTCGACAGCTGACTTTGGTG 1500
Db 481 GlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTGGACTGGCGCGAGCCGCTAGCCGATCGCTTGACTGCTCCAAAGTCAACCTC 1560
Db 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
QY 1561 GACGGCGGCGCCCTCCGACTGTTGAGCAATATTCAGACATCTCTTGCTGCTCCCGCTT 1620
Db 521 AspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGCAAGCTCTCTTTTGGAGCGCGGCACAAAGCAGGTTATCTTATTAATTAT 1680
Db 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyr 560
QY 1681 AATACTACTGCTAGTGACAGATTCTGATTGAAATGCTGCGGCGCATCGGTCGCCATT 1740
Db 561 AsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580
QY 1741 TCACCTATACACAGCTTGGGCGCGTCCGGTCGCAATTCCTGCGCGCGCGGTTTGTG 1800
Db 581 SerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeu 600
QY 1801 GCTCCAGCTCGGCGCTGCTGCTGAGGAGTACTTTTGATATTCGCGGCGCGCGGCAC 1860
Db 601 AlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHis 620
QY 1861 ACATTGATGACTTCTGCGCTGAATGCGCGCTTTAGGCCTCCAGGTTGTGCTTTCCAG 1920
Db 621 ThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGln 640
QY 1921 TCACCTGCTGCTAGCTCCAGCGCTTAAAGTTAAAGTGGGTAAACCTCGGAGTTG 1977
Db 641 SerThrValAlaGluGlnArgLeuLysValLysValLysThrArgGluLeu 659
```

RESULT 3

AAR39308
ID AAR39308 standard; protein; 659 AA.
XX
AC AAR39308;
XX
DT 25-MAR-2003 (revised)
DT 14-FEB-1994 (first entry)
XX Mexico strain HEV ORF2 putative virus capsid protein.
DE
XX Hepatitis E virus; vaccine; neutralising antibodies; infection; block;
KW open reading frame; antibodies.
XX
OS Hepatitis E virus.
XX
PN W09314208-A2.
XX
PD 22-JUL-1993.
XX
PF 19-JAN-1993; 93WO-US000475.
XX
PR 17-JAN-1992; 92US-00822335.
PR 20-APR-1992; 92US-00870985.
XX

(GENE-) GENELABS TECHNOLOGIES INC.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Reyes GR, Bradley DW, Twu J, Purdy MA, Tam AW, Krawczynski KZ;
XX
DR WPI; 1993-243223/30.
DR N-PSDB; AAQ46814.
XX

PT Antigen and antibody vaccines against hepatitis E virus infection -
PT contain peptide(s) derived from capsid protein C-terminal or antibodies
against protein.

XX Disclosure; Fig 7; 43pp; English.

XX The sequence is that of the putative virus capsid protein encoded by
CC Mexico strain hepatitis E virus (HEV) open reading frame ORF2. This
CC protein or peptide fragments of it may be used in a vaccine composition
CC for immunising an individual against HEV. Antibodies raised against these
CC peptides can also be used in such vaccines. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX

SQ Sequence 659 AA;

Alignment Scores:

Pred. NO.: 0 Length: 659
Score: 572.00 Matches: 658
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 86.80% Indels: 2
DB: Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR39308 (1-659)

QY 1 ATGCGCCCTAGCCCTTTTCTGCTGTTGTTCTCTTGTCTGCTTATGTCCTGTCGCGCA 60
Db 1 MetArgProArgProLeuLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CCGACCGGTACGCGCTCTGCGCGCGCTGCTGCGCGCGCGAGCGCGGTACCGGCGGTG 120
Db 21 ProThrGlyGlnProSerGlyArgArgGlyArgGlyArgSerGlyGlyThrGlyGlyGly 40
QY 121 TCTCGGGGTAGCCGGTGTGATTCAGCCCTTCGCAATCCCTATATTCATCCAAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTTGCGCCAGACGTTGCGGTCGCGTCCGCTTCGACCTCGCTTCGCCAACACGACC 240
Db 61 ProPheAlaProAspValAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80

QY 241 CGGCCACTTGCTCCACTTGGCGAGATCAGCCGCGCCCTCCGCTGCTCCCGTCCG 300
 Db |||||
 QY 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
 Db |||||
 QY 301 CGACCTGCCACAGCCGGGCTGGCGGCTGACGGCTGTGGCGCTGCCCATGACACTCA 360
 Db |||||
 QY 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
 Db |||||
 QY 361 CCGCTCCGAGCCTGATCTCCGGGTGGGAAATCTACGGCGCCAGTATATTTGTCTACT 420
 Db |||||
 QY 121 ProValProAspValAspSerArgGlyAlaAlaLeuArgGlnTrpAsnLeuSerThr 140
 Db |||||
 QY 421 TCACCCCTGACATCTCTGCGCTCGGCACATAATTTAGTCTGTATGACGCCGCCCT 480
 Db |||||
 QY 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuThrAlaAlaProLeu 160
 Db |||||
 QY 481 AATCCCGCTCTGCGCTGACGAGCGTACTTAATCACTCACATTAATGGCCACAGAGGCTCC 540
 Db |||||
 QY 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
 Db |||||
 QY 541 AATATGACAGTACCGGTTGCCGGCTACTATCCGTTACCGGCCCTAGTGCCTAAT 600
 Db |||||
 QY 181 AsnTrpAlaGlnTrpArgValAlaArgAlaThrIleArgTrpArgProLeuValProAsn 200
 Db |||||
 QY 601 GCAGTTGGAGGCTATGCTATATCAATTTCTTCTGGCTCAAAACAACCAACCCCTACA 660
 Db |||||
 QY 201 AlaValGlyGlyTrpAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
 Db |||||
 QY 661 TCTGTTGACATGAATTCCTACTTCCACTGATGTCAGGATTCCTGTCACCTGGCATA 720
 Db |||||
 QY 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
 Db |||||
 QY 721 GCATCTGAATGTGATCCCAAGGAGCGCTTCACTACCGCATCAAGTTGGCGCTCG 780
 Db |||||
 QY 241 AlaSerGluLeuValIleProSerGluArgLeuHisTrpArgAsnGlnGlyTrpArgSer 260
 Db |||||
 QY 781 GTTGAGACATCTGTGTTGTGAGAGAGAGCCACTCCGCTCTGTGTCATGTTATGCATA 840
 Db |||||
 QY 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
 Db |||||
 QY 841 CATGGCTCTCCAGTTAACTCTATACCAATAACCCCTTATACCGGTGCCCTTGGCTTACTG 900
 Db |||||
 QY 281 HisGlySerProValAsnSerTrpThrAsnThrProTrpThrGlyAlaLeuGlyLeuLeu 300
 Db |||||
 QY 901 GACTTTCGCTTAGAGTTGATTCGCAATCTACCACTGTACCAACCAATACAGTGTG 960
 Db |||||
 QY 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrTrpCysAsnThrAsnThrArgVal 320
 Db |||||
 QY 961 TCCGTTTACTCCAGCACTGCTGCTCACTCCGCGGAGGGCGGAGCTGCGGAGCTG 1020
 Db |||||
 QY 321 SerArgTrpSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
 Db |||||
 QY 1021 ACCAACAATGACGACCCAGGTTTCATGAAGATCTCCACTTACCGGCTTAAATGGGTA 1080
 Db |||||
 QY 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
 Db |||||
 QY 1081 GGTGAAGTCCGCGGAGTAGCTTCAATTAACCTTACTTAACCTGCTGACACGCTCTCCG 1140
 Db |||||
 QY 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
 Db |||||
 QY 1141 GGGCTCCGACAGAAATTAATTTCTGCTGCGCTGCGGGCACTGTTTTATTCCCGCCGGT 1200
 Db |||||
 QY 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTrpSerArgProVal 400
 Db |||||
 QY 1201 GTCTCAGCAATGGCGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAATCTCAGCAG 1260
 Db |||||
 QY 401 ValSerAlaAsnGlyGluProThrValLysLeuTrpThrSerValGluAsnAlaGlnGln 420
 Db |||||
 QY 1261 GATTAAGGTGTGCTATCCCCCAGCATATCGATCTTGATTCGGTGTGCTGCTATTTCAG 1320
 Db |||||
 QY 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValIleGln 440
 Db |||||
 QY 1321 GATTATGACACCATGACGAGATCGGCCACCCCGCTCGCTCGCCATCTCGGCT 1380
 Db |||||

Db 441 AspTrpAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
 QY |||||
 QY 1381 TTTTCTGTTCTCCGAGCAAAATGATGTACTTGGCTGCTCCCTCACTGCGCCGAGTATGAC 1440
 Db |||||
 QY 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTrpAsp 480
 Db |||||
 QY 1441 CAGTCCACTTACCGGTCTGCAACTGGCCCGGTTTATATCTCGGACAGCGTACTTTGGTG 1500
 Db |||||
 QY 481 GlnSerThrTrpGlySerSerThrGlyProValTrpIleSerAspSerValThrLeuVal 500
 Db |||||
 QY 1501 AATGTTCCGACTCGCGCGAGCGCGTAGCCGATCGCTTCACTGCTCCAAAGTC-ACCCT 1559
 Db |||||
 QY 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSer-LysLeuThrLe 520
 Db |||||
 QY 1560 CGAGCGGCGCCCTCCGACTCTTGGCAATATTCGAACACATCTTCTGCTGCTCCCT 1619
 Db |||||
 QY 520 uAspGlyArgProLeuProThrValGluGlnTrpSerLysThrPhePheValLeuProLe 540
 Db |||||
 QY 1620 TCGTGGCAAGCTCTCCCTTTGGGAGCGCGCACAAACAAAGCAGGTTATCTTATATTA 1679
 Db |||||
 QY 540 uArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTrpTrpTrpAsnTy 560
 Db |||||
 QY 1680 TAATACTACTGCTAGTACCCAGATTCGATGTAATAATGCTCCGCGCATCGGTGCGCAT 1739
 Db |||||
 QY 560 rAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580
 Db |||||
 QY 1740 TTCAACCTATACACAGGCTTGGGCGCTGCGGTGCGCATTTCTCGCGCGCGGTTT 1799
 Db |||||
 QY 580 eSerThrTrpThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLe 600
 Db |||||
 QY 1800 GGCTCCACGCTCCGCGCTGCTGCTGAGAGGATCTTTTATTCGCGCGCGGCGCA 1859
 Db |||||
 QY 600 uAlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTrpProGlyArgAlaHi 620
 Db |||||
 QY 1860 CACATTTGATGACTTCTGCCCTGAATCCGCGCTTTAGGCTCCAGGTTGTGCTTTCCA 1919
 Db |||||
 QY 620 sThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGl 640
 QY 1920 GTCACTGCTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGTAAACTCGGAGTTG 1977
 Db |||||
 QY 640 nSerThrValAlaGluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 659
 Db |||||
 RESULT 4
 AAW93387
 ID AAW93387 standard; protein; 659 AA.
 XX
 AC AAW93387;
 XX
 DT 11-JUN-1999 (first entry)
 XX
 DE Human HEV ORF 2 protein from strain Mexico.
 XX
 KW Swine hepatitis E virus; HEV; cross-reaction; antibody; human; therapy;
 XX vaccine; immunise; infection; detection; diagnosis; prevention.
 XX
 OS Hepatitis E virus.
 XX
 PN WO9904029-A2.
 XX
 PD 28-JAN-1999.
 XX
 PF 17-JUL-1998; 98WO-US014665.
 XX
 XX 18-JUL-1997; 97US-0053069P.
 PR
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Meng X, Emerson SU, Purcell RH;
 XX
 DR WPI; 1999-132270/11.
 XX
 PT New isolated swine hepatitis E virus - used to develop products for the

PT diagnosis, prevention and treatment of hepatitis E virus infection in
XX mammals, particularly humans.

Example 1; Fig 3A; 70pp; English.

CC This invention describes a swine hepatitis E virus (HEV) and its natural
CC mutants which are capable of cross-reacting with antibodies reactive with
CC a human HEV strain or natural mutants. The HEV and the proteins can be
CC used in vaccines for immunising against HEV infection. The swine HEV can
CC be used in humans to prevent possible infection by human HEV. The swine
CC HEV can also be used as a therapeutic treatment for infection by other
CC strains of HEV. The swine HEV can also be used for the production of
CC antibodies which can be used in therapy, detection and diagnosis. The
CC products can also be used for determining the susceptibility of cells or
CC organs to infection with swine HEV. The swine HEV is particularly useful
CC for the development of agents for the prevention, treatment and detection
CC of human HEV because it is not a human virus and thus can be handled both
CC experimentally and clinically without fear of severe infection and/or
CC contamination

SQ Sequence 659 AA;

Alignment Scores:

Pred. No.: 0 Length: 659
Score: 558.00 Matches: 658
Percent Similarity: 99.85% Conservative: 0
Best Local Similarity: 99.85% Mismatches: 1
Query Match: 84.67% Indels: 0
DB: 2 Gaps: 0

US-09-851-410a-10_COPY_5117_7096 (1-1980) x AAW93387 (1-659)

QY 1 ATGCGCCCTAGCCCTCTTTGCTGTTGTTCTCTGCTTCTGCTATGTTGCCGCGCA 60
DB 1 MetArgProArgProLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CGACCGGTACCGCTCTGCGCGCGCTGCTGGCGCGCGAGCGCGGTACCGCGGTGT 120
DB 21 ProThrGlyGlnProSerGlyArgArgArgGlyArgArgSerGlyGlyThrGlyGly 40
QY 121 TTCTGGGTGACCGGTGATCTCAGCCCTTCGCAATCCCTATATTCATCCCAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTTGCCAGACGTTGCTGCCGCTGCTCGCGGTCTGGACCTTCGCGCAACAGCC 240
DB 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGGCCACTTGCTCCACTTGGCGAGATCAGGCCGCGCGCGCTCCGCTGCTCCGCTGCG 300
DB 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArg 100
QY 301 CGACCTGCCACAGCGCGCTGCGCGCTGACGGCTGTGGCGCTTCGCCATGACACCTCA 360
DB 101 ArgProThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCGCTCCCGAGCTGATCTCGCGGTGCAATTCACCGCGCGAGTATAATTTGCTACT 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCACCCCTGACATCTCTGCGCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 480
DB 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AATCGCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 540
DB 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATATGACAGTACCGGTGCGCGCTACTATCCGTTACCGGCCCTAGTGCCTAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaAlaAlaAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCAGTTGGAGGCTATGCTATATCCATTCTTCTGCGCTCAACACCAACCCCTTACA 660

DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCTGTTGACATGAATCCATTAATCTCCACTGATGTCAGGATCTTCTTCAACCTGCGATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATTCGTTCATCCCAAGCAGCGCTTCACTACCGCAATCAAGCTGGCGCTCG 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTGAGACATCTGCTGCTGCTGAGGAGGAGCCCTCCGCTCTTCTCATGTATATGATA 840
DB 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCTCCAGTAACTCTTATACCAATACCCCTTATACCGGTGCGCTTGGCTACTG 900
DB 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAAACACCAATACACGTGTG 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTACTCCAGCACTGCTCTCATCTCCGCGCGAGGCGCGAGCTCGGAGCTG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 AGCACAACTGCGAGCCAGCTTCATGAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080
DB 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAAGTCGCGCGCGGATAGTCTTAACATTACTTAACCTTGTGACACGCTCTCGCG 1140
DB 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGCTCCCGACAGAAATTAATTTGCTGCGCTGCGCGGCAACTGTTTATTCGCGCGGFT 1200
DB 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GCTCAGCAATGCGGAGCCACCACTGAGCTCTATATCATCATGTCGAGATGCTCAGCAG 1260
DB 401 ValSerAlaAsnGlyGluProThrValIleLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAGGCTGTTGCTATATCCCGATATCCCGATATCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 421 AspLysGlyValAlaIleProHisAspLeuLeuAspLeuGlyAspSerArgValIleGln 440
QY 1321 GATTATGACACCGAGCATGAGGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1380
DB 441 AspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTTCTCCGAGCAATGATGATGCTTGGCTGCTCCCTCAGTGCAGCGAGTATGAC 1440
DB 461 PheSerValLeuArgAlaAsnAspValLeuLeuTrpLeuSerLeuThrAlaAlaGlnTyrAsp 480
QY 1441 CAGTCCATTTACGGCTGCTCAACTGCGCGCGGTTTATATCTCGGACAGCGTGAATTTG 1500
DB 481 GlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTGCGACTGCT 1560
DB 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
QY 1561 GACGGCGCGCGCTCCCGACTGTTGAGCAATATTCAGAGATCTTGTGCTCCCGCTT 1620
DB 521 AspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGGCAAGCTCTCTTTTGGAGCGCGCGCACAAAGAGAGGTATTCCTTTAATAATTAT 1680
DB 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyr 560
QY 1681 AATACTACTGCTAGTGACAGATTCGTGATTAATAATGCTGCGCGCGCTTCCCGCTT 1740


```
QY 1017 GGTGACCACTGACGACCAAGGTTTCATGAAGATCTCCACTTTACCGGCTTAATGG 1076
Db |||
QY 340 uLeuThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGI 360
Db |||
QY 1077 GGTAGTGAAGTCGGCGCGGGATAGCTTAACTTAACTTAACTTAACTTAACTTAACTT 1136
Db |||
QY 360 yValGlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLe 380
Db |||
QY 1137 CGCGCGGCTCCGACAGATTAAATTCGTCGGCTGGCGGCACTCTTTTATTCGCGCC 1196
Db |||
QY 380 uGlyGlyLeuProThrGluLeuLeuSerSerAlaGlyGlyGlnLeuPheTySerArgPr 400
Db |||
QY 1197 GGTGTCTCAGCAATGGGAGCAACCGTGAAGCTCTATACATCAGTGGAGAATGCTCA 1256
Db |||
QY 400 oValValSerAlaAsnGlyGluProThrValLysLeuTyThrSerValGluAsnAlaGI 420
Db |||
QY 1257 GAGGATAGGGTGTGCTATCCGCCACGATATCGATCTTGGTATCGCGTGTGTCAT 1316
Db |||
QY 420 nGlnAspLysGlyValAlaIleProHisAspLeuGlyAspSerArgValValI 440
Db |||
QY 1317 TCAGGATTATGACCAACGACATGAGCAGGATCGGCCACCGCTCGCTCGGCATCTCG 1376
Db |||
QY 440 eGlnAspTyAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerAr 460
Db |||
QY 1377 GCCTTTTCTGTCTCCGAGCAATGATGATCTTTGGCTGTCTCCTCACTGCGAGCCGAGTA 1436
Db |||
QY 460 gProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTy 480
Db |||
QY 1437 TCACGATCCACTTACGGTGTCTCAACTGCGCGGTTTATATCTCGGACAGGTGACTTT 1496
Db |||
QY 480 rAspGlnSerThrTyGlySerSerThrGlyProValTyIleSerAspSerValThrLe 500
Db |||
QY 1497 GGTGAATGTGTGCTGGCGCGCAGCGCTAGCCGATCGCTTGACTGTGCTCAAAATGCAC 1556
Db |||
QY 500 uValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValTh 520
Db |||
QY 1557 CTGCGACGGCGGCGGCTCCGACTGTGTGAGCAATATTCGAACACATCTTTGTGCTCC 1616
Db |||
QY 520 rLeuAspGlyArgProLeuProThrValGluGlnTySerLysThrPhePheValLeuPr 540
Db |||
QY 1617 CCTTCGTGGCAGCTCTCTTTGGAGCGCGGCACAAAGAGGTTATCTTATTA 1676
Db |||
QY 540 oleuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyProTyAs 560
Db |||
QY 1677 TTATATATCTACTGTAGTACAGATTCTGATTGAATGCTGCGGCCATCGGTGCGC 1736
Db |||
QY 560 nTyraThrThrAlaSerAspGlnIleLeulleGluAsnAlaAlaGlyHisArgValAl 580
Db |||
QY 1737 CATTTCAACCTATACACAGGCTTGGGCGCGGTCCGGTCGCCATTTCTGCGCGCGGT 1796
Db |||
QY 580 aileSerThrTyThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaVa 600
Db |||
QY 1797 TTTGGCTCCAGCTCGGCTGCTGCTGAGGATCTTTGATTATCCGGGGGGGC 1856
Db |||
QY 600 lleuAlaProArgSerAlaLeuAlaLeuGluAspThrPheAspTyProGlyArgAl 620
Db |||
QY 1857 GCACACATTTGATGATCTTGGCCCTGAATGCGCGCTTTAGCGCTTCCAGGGTTGTGCTTT 1916
Db |||
QY 620 aHisThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPh 640
Db |||
QY 1917 CAGTCACTGTGCTGAGCTCCAGCGCTTAAAGTTAAGTGGGTAAACTCGGAGTT 1976
Db |||
QY 640 eGlnSerThrValAlaGluLeuGlnArgLeuLysValGlyLysThrArgGluLe 660
Db |||
QY 1977 G 1977
Db 660 u 660

RESULT 6
ID AAR96092
XX AAR96092 standard; protein; 549 AA.
AC AAR96092;
```

```
XX 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX
XX Hepatitis E virus (Mexico strain) 62K antigen.
XX
XX HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
KW capsid.
XX
XX Hepatitis E virus; Mexico strain.
XX
XX WO9612807-A2.
XX
XX 02-MAY-1996.
XX
XX 23-OCT-1995; 95WO-US013703.
XX
XX 24-OCT-1994; 94US-00327952.
XX
XX 13-OCT-1995; 95US-00542634.
XX
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX
XX Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
XX
XX WPI; 1996-230608/23.
XX
XX N-PSDB; AAT27110.
XX
XX Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
PT diagnostic reagents for determining HEV infection and in vaccines.
XX
XX Claim 4; Page 84-86; 125pp; English.
XX
XX 62K antigen (AAR96092) comprises the C-terminal 549 amino acids of
CC Hepatitis E virus (HEV) Mexico strain capsid protein (see also AAR96090).
CC It is obtd. by PCR amplification of HEV Mexico ORF-2 (AAT27108),
CC insertion of amplified DNA (AAT27110) into vector pGEX and expression in
CC E. coli cells. Expression of full-length ORF-2 in Sf9 insect cells using
CC a baculovirus system also results in prodn. of 62K antigen (see also
CC AAR96102 and AAR96104). 62K represents an improved antigen, in comparison
CC to bacterial expressed proteins, for use in HEV diagnostic assays, and
CC also has excellent immunogenic properties. (Updated on 16-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 549 AA;

Alignment Scores:
Pred. No.: 0 Length: 549
Score: 401.00 Matches: 547
Percent Similarity: 99.45% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 60.85% Indels: 3
DB: 2 Gaps: 1

US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR96092 (1-549)

QY 334 GCTGTGGCGCGCTGCGCATGACACCTCCACCGTCCGAGCTTGATCTCGCGTCAATT 393
Db |||
Db 1 AlaValAlaProAlaHisAspThrSerProValProAspValAspSerArgGlyAlaIle 20
QY 394 CTCACGCCGCCAGTATAATTTGTCTACTTCCACCCCTGACATCCTCTGTGGCTCTGGCACT 453
Db |||
Db 21 LeuArgArgGlnTyraAsnLeuSerThrSerProLeuThrSerSerValAlaSerGlyThr 40
QY 454 AATTAGTCTCTGATGCGCGCCCTTAAATCCGCTCTGCGCTCGCAGACGCTACTAAT 513
Db |||
Db 41 AsnLeuValLeuTyraAlaProLeuAsnProProLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACTCATATTATGGCCACAGAGCGCTCCCAATTATGACAGTACCGGGTTGCCCGCGCTACT 573
Db |||
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyraGlnTyraArgValAlaAlaThr 80
QY 574 ATCCGTTACCGGCCCTAGTGCCTAATGCAGTTGGAGGCTATGTCTATATCCATTTCTTTC 633
```

Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
Qy 634 TGGCTCAACCAACCAACCCCTACATCTGTGTGACATGAATCCATTACTTCCACTGAT 693
Db 101 TrpProGlnThrThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120
Qy 694 GTGAGGATCTTGTTCACCTGGCATAGCATCTGAATGTGCATCCCAAGCGAGCGCTT 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
Qy 754 CACTACCGCAATCAAGGTGGCGTTCGCTGAGACATCTGGTGTGTGTGAGGAGGAAGCC 813
Db 141 HisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGluGluAla 160
Qy 814 ACCTCCGGTCTTGTTCATGTTATGATACATAGCTCTCCAGTTAACTCTTATACCAATACC 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThr 180
Qy 874 CCTTATACCGGTGGCTTGTACTCGACTTTCCTTAGAGCTTGAGTTTCGCAATCTC 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsnLeu 200
Qy 934 ACCACCTGTAAACACCAATACACGTGTGTCCTGTTACTCCAGCACTGCTCGTCA----CTC 989
Db 201 ThrThrCysAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
Qy 990 CGCCGAGGGCCGACGGAGCTCGGGAGCTGACCACTGACCACTGACGACCGAGTTTCATGAA 1049
Db 221 Arg-ArgGlyAlaAspGlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetLy 240
Qy 1050 AGATCTCCACTTACCGCCCTTAATGGGTAGGTGAAGTCCGCGCGGATAGTCTTAAC 1109
Db 240 sAspLeuHisPheThrGlyLeuAsnGlyValGlyLeuValGlyArgGlyIleAlaLeuTh 260
Qy 1110 ATTACTTAACCTGTGACAGCTCCTCGCGGGCTCCCGAGAGATTAATTAATTCGTCCGC 1169
Db 260 rLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAl 280
Qy 1170 TGGCGGGCACTGTTTTATCCCGCCCGTGTCTACGCCAATGGCGAGCAACCGTGAA 1229
Db 280 aGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrVally 300
Qy 1230 GCTTATACATCAGTGGAGATGCTCAGCAGATAAGGTGTGCTATCCCGCCAGTAT 1289
Db 300 sLeuTyrThrSerValGluAsnAlaGlnAspLysGlyValAlaIleProHisAspIl 320
Qy 1290 CGATCTTGGTCAATTCGCGTGTGTCATTTCAGGATTTATGACCAACACCATGACGAGATCG 1349
Db 320 eAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspAr 340
Qy 1350 GCCACCCCGTCGCTGGCGCATCTCGGCCTTTTCTGTTCTCGAGCAATGATGTA 1409
Db 340 gProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValle 360
Qy 1410 TTGCTGTCCCTCACTGACGCGAGTANGACAGTCCACTTACGGTGTGTCAGTCCGCGCC 1469
Db 360 uTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPr 380
Qy 1470 GGTATTATCTCGGACAGCGTGACATTTGGTGAATGTTGCGATGCGCGCGCGCGTACG 1529
Db 380 oValTyrIleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAl 400
Qy 1530 CCGATCGTGTGATCGGTCCAAAGTCACCCCTGACGCGGCGGCCCTCCGACTGTTGAGCA 1589
Db 400 aArgSerLeuAspTrpSerLysValThrLeuAspGlyArgProLeuProThrValGluGl 420
Qy 1590 ATATTCCAGACATTTGTGCTCCCTTTCGTCGAGCTCTCTTTTGGGAGCGCG 1649
Db 420 nTyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGl 440
Qy 1650 CACAAACAAAGCAGGTATCTCTATAATTAATACTACTGCTAGTACCAAGATTCTGAT 1709

Db 440 yThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeuIl 460
Qy 1710 TGAATAATGTCGGCCCATCGGTCCGCAATTCACCTATACACACGAGCTTGGGCGCG 1769
Db 460 eGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrArgLeuGlyAlaGl 480
Qy 1770 TCCGGTCCCATTTCTCGCGCGCGGTGGTTCACCGTCCCGCTCCCGCTGCTGCTGGA 1829
Db 480 yProValAlaIleSerAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuGl 500
Qy 1830 GGATACATTTGATATATCCGGCGCGCGCACACATTTGATGATCTTCCCTGATGCGG 1889
Db 500 uAspThrPheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysAr 520
Qy 1990 CGCTTTAGGCTCCAGGTTGTGCTTTCAGTCAACTGCTGAGCTCCAGCGCCTTAA 1949
Db 520 gAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLy 540
Qy 1950 AGTTAAGTGGGTAAAACTCGGAGTTG 1977
Db 540 sValLysValGlyLysThrArgGluLeu 549
RESULT 7
AAR96102
ID AAR96102 standard; protein; 540 AA.
XX AC AAR96102;
XX 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX Hepatitis E virus (Mexico strain) recombinant 62K antigen.
DE HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
XX diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
KW capsid.
XX Hepatitis E virus; Mexico strain.
OS WO9612807-A2.
FN 02-MAY-1996.
XX 23-OCT-1995; 95WO-US013703.
XX 24-OCT-1994; 94US-00327952.
PR 13-OCT-1995; 95US-00542634.
XX (GENE-) GENELABS TECHNOLOGIES INC.
PI Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
XX WPI; 1996-230608/23.
DR Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
PT diagnostic reagents for determining HEV infection and in vaccines.
XX Claim 6; Page 93-95; 125pp; English.
XX DNA (AAT27110) coding for the 62K antigen (AAR96092) of hepatitis E virus
(HEV) Mexico strain capsid protein was cloned into baculovirus expression
vector pBacIII and recombinant 62K was expressed in Sf9 insect cells.
CC High levels of expression were obtd. and the recombinant 62K was obtd. in
CC over 95% purity. However, C-terminal processing resulted in the deletion
CC of 9 or 23 amino acids from 62K, giving 2 related polypeptide species
CC (AAR96102 and AAR96104). Similar results were obtd. with HEV Burma 62K
CC antigen (see also AAR96101 and AAR96103). Recombinant 62K represents an
CC improved antigen, in comparison to bacterial expressed proteins, for use
CC in HEV diagnostic assays, and also has excellent immunogenic properties.
XX (Updated on 16-OCT-2003 to standardise OS field)
SQ Sequence 540 AA;

Alignment Scores:

Pred. No.: 0 Length: 540
Score: 392.00 Matches: 538
Percent Similarity: 99.45% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 59.48% Indels: 3
DB: 2 Gaps: 1

US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR96102 (1-540)

QY 334 GCTGTGGCGCTGCCCATGACACTCACCGTCCCGGACGTTGAATCTCGCGTGCAT 393
Db 1 AlaValAlaProAlaHisAspThrSerProValProAspValAspSerArgGlyAlaIle 20
QY 394 CTACGGCGCCAGTAAATTTGCTACTTCCACCCCTGACATCCTCTGCGCTCGCACT 453
Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaSerGlyThr 40
QY 454 AATTTAGTCTGTATGACGCCCCCTTAATCCGCTCTGCGCTGACGAGCGTACTAAT 513
Db 41 AsnLeuValLeuTyrAlaAlaProLeuAsnProProLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACTCACATTTATGCCACAGAGGCTCCAATTATGACAGTACCGGTTCCCGCGCTACT 573
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaArgAlaThr 80
QY 574 ATCCGTTACCGGCCCTAGTGCCTAATGCAGTTGGAGGCTATGCTATATCCATTTCTTC 633
Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCTCAAAACACCAACCCCTACATCTGTTGACATGAATTCATTTCCACTGAT 693
Db 101 TrpProGlnThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120
QY 694 GTCAGATCTTGTTCACCTGGCATAGCATCTGAATTTGGTATCCCAAGACGAGCCCTT 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTACCCCAATCAAGGTTGGCGCTCGGTTGAGACATCTGGTGTGCTGAGGAGGAGCC 813
Db 141 HisTyrArgAsnGlnGlyTrpArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCCGGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873
Db 161 ThrSerGlyLeuValMetLeuGlyHisGlySerProValAsnSerTyrThrAsnThr 180
QY 874 CCTTATACCGTGGCTTGGCTTATGAGTCTGCTTGGCTTGGCTTGGCTTGGCTTGGCT 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsnLeu 200
QY 934 ACCACTGTAAACCAATACAGTGTGCTCCGCTTACTCCAGCACTGCTGCTCA---CTC 989
Db 201 ThrThrCysAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
QY 990 GCGCGAGGGCGGACGGAGCTGCGGAGCTGACCAACACTGACGACCAAGTTCATGAA 1049
Db 221 Arg-ArgGlyAlaAspGlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetIy 240
QY 1050 AGATCTCCACTTTACCGGCTTAAATGGGTAGTGAAGTCCGCGCGGATAGCTCAAC 1109
Db 240 saspleuHisPheThrGlyLeuAsnGlyValGlyGluValGlyAlaLeuThr 260
QY 1110 ATTACTTAACCTTGTGACAGCTCTCTCGCGGGCTCCCGACAGAAATTAATTTCTCGCG 1169
Db 260 rLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAl 280
QY 1170 TGGCGGCAACTGTTTATTCGCGCGGTTGCTCAGCCAAATGGCGAGCCAAACGCGTAA 1229
Db 280 adGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrVally 300
QY 1230 GCTCTATACATCAGTGGAGATGCTCAGCAGGATAAGGTGTTGCTATCCCGCCAGATAT 1289
Db 300 sLeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyValAlaIleProHisAspI 320

QY 1290 CGATCTTGTGATTCGCTGTGCTCATTGAGGATTTAGACACACGATGAGCAGATCG 1349
Db 320 eAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspAr 340
QY 1350 GCCACCCCGTCCGCTGCGCCATCTCGGCTTTTCTGTTCTCCGAGCAAAATGATGTACT 1409
Db 340 gProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValle 360
QY 1410 TTGGCTGTCCCTCACCTGACGCGAGTATGACAGTCCACTTACGGGTCTGTCACCTGCC 1469
Db 360 uTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPr 380
QY 1470 GGTATTATCTCGGACAGCTGACTTGTGTAATGTTGGACTGGCGCGCAGCCCTAGC 1529
Db 380 oValTyrIleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAl 400
QY 1530 CCGATCGCTTGATCGTCCAAAGTCACCTCCAGCGCGCGCCCTCCCGACTGTTTCAGCA 1589
Db 400 aArgSerLeuAspTrpSerLysValThrLeuAspGlyArgProLeuProThrValGlu 420
QY 1590 ATATTCCAAGACATCTTGTGCTCCCTTCTGTCGCAAGCTCTCTTTTGGAGGCCGG 1649
Db 420 nTyrSerTyrThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaG 440
QY 1650 CACAACAAAGCAGGTTATCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1709
Db 440 yThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeu 460
QY 1710 TGAATATGTCGCGGCTCGGCTCGCCATTTCACTATACCACTATACCACTGTTGGGCGCG 1769
Db 460 eGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrArgLeuGlyAlaG 480
QY 1770 TCGGTCGCAATTCGCGGCGCGGTTTGGCTCCACCTCGGCTCGGCTCGGCTCGGCTCG 1829
Db 480 yProValAlaIleSerAlaAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuLeu 500
QY 1830 GGATACTTTTGTATTCGCGGCGCGGCGCACACATTTGATGACTTCTGCTGCTGCTGCTG 1889
Db 500 uAspThrPheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysAr 520
QY 1890 CGCTTAGCCTCCAGGTTGTCTTCCAGTCAACTGTCGCTGAGCTCCAGCGCTTAA 1949
Db 520 gAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeu 540
QY 1950 A 1950
Db 540 S 540
RESULT 8
AAR96104
ID AAR96104 standard; protein; 525 AA.
XX
AC AAR96104;
XX
DT 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX
DE Hepatitis E virus (Mexico strain) recombinant 62K antigen.
XX
KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;
XX capsid.
OS Hepatitis E virus; Mexico strain.
XX
PN WO9612807-A2.
XX
PD 02-MAY-1996.
XX
PF 23-OCT-1995; 95WO-US013703.
XX
PR 24-OCT-1994; 94US-00327952.

PR 13-OCT-1995; 95US-00542634.
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX PI Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
XX DR MPI; 1996-230608/23.
XX PS Claim 8; Page 97-98; 125pp; English.
XX
XX DNA (AAT27110) coding for the 62K antigen (AAR96092) of hepatitis E virus
CC (HEV) Mexico strain capsid protein was cloned into baculovirus expression
CC vector pBluscript and recombinant 62K was expressed in Sf9 insect cells.
CC High levels of expression were obtd. and the recombinant 62K was obtd. in
CC over 95% purity. However, C-terminal processing resulted in the deletion
CC of 9 or 23 amino acids from 62K, giving 2 related polypeptide species
CC (AAR96102 and AAR96104). Similar results were obtd. with HEV Burma 62K
CC antigen (see also AAR96101 and AAR96103). Recombinant 62K represents an
CC improved antigen, in comparison to bacterial expressed proteins, for use
CC in HEV diagnostic assays, and also has excellent immunogenic properties.
CC (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 525 AA;
SQ

Alignment Scores:
Pred. No.: 0 Length: 525
Score: 377.00 Matches: 523
Percent Similarity: 99.43% Conservative: 0
Best Local Similarity: 99.43% Mismatches: 1
Query Match: 57.21% Indels: 3
DB: 2 Gaps: 1

US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR96104 (1-525)

QY 334 GCTGTGGCGCTGCCATGACCTACCTCCGTCGGAGCTGTGATCTCCGGTGCAT 393
Db 1 AlaValAlaProAlaHisAspThrSerProValProAspValAspSerA-gglyAlaIle 20

QY 394 CTAGCGCGCCAGTAAATTTGTCTACTTCACCCCTGCATCCTCTGTGGCTCTGGCACT 453
Db 21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaSerGlyThr 40

QY 454 AATTAGTCTGTATGACGCCCCCTTAATCCGGCTCTGGCGCTGACGAGCGTACTAAT 513
Db 41 AsnLeuValLeuTyrAlaAlaProLeuAsnProProLeuProLeuGlnAspGlyThrAsn 60

QY 514 ACTCACATTATGCCACAGAGGCTCCAATTATGCACAGTACCGGGTTGCCGCTACT 573
Db 61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaAla-gAlaThr 80

QY 574 ATCCGTTTACCGGCCCTTAGTGCCTTAATGCAGTTGGAGGCTATGCTATATCCATTCTTTC 633
Db 81 IleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSerIleSerPhe 100

QY 634 TGGCCTCAACACACCAACCCCTACATCTGTTGACATGAATTCCTTCCACTGAT 693
Db 101 TrpProGlnThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120

QY 694 GTCAGGATTCTGTTCACCTGGCATAGCATCTGAATTGGTTCATCCCAAGCGAGCGCTT 753
Db 121 ValArgIleLeuValGlnProGlyIleAlaSerGlnLeuValIleProSerGluArgLeu 140

QY 754 CACTACCGCAATCAAGGTTGGCGCTCGGTTGAGACATCTGTTGTGCTGAGGAGGAAGCC 813
Db 141 HisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGluGluGluAla 160

QY 814 ACCTCCGGTCTGTGTCATGTTATGCATACATGGCTCTCCAGTTAACTCCTATACCAATACC 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThr 180

RESULT 9

QY 874 CTTTATACCGGTGCCCTTGGCTTACTGCAGCTTTGCCCTTAGAGCTTGAGTTTCGCAATCTC 933
Db 181 ProTyrThrGlyAlaLeuLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200

QY 934 ACCACCTGTAAACACCAATACACCGTGTGTCCCGTTACTCCAGCACTGCTGCTCA----CTC 989
Db 201 ThrThrCysAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220

QY 990 GCGCCGAGGGCGGACCGGAGCTCGGAGCTGACCACTGCAGCCACCACTGAGTTTCATGAA 1049
Db 221 Arg-ArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLeu 240

QY 1050 AGATCTCCACTTACCGGCTTAATGGGTAGGTGAAGTGGCGCGCGGATAGCTCTAATC 1109
Db 240 SASpLeuHisPheThrGlyLeuAsnGlyValGlyGluValGlyArgGlyIleAlaLeu 260

QY 1110 ATTACTTTAACTTGTCTGACACCGTCTCCGCGGGCTCCCGACAGAAATTAATTTCTCGCGC 1169
Db 260 rLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAl 280

QY 1170 TGGCGGGCAACTGTTTATTCCCGCCCGGTTGTCTCAGCCAAATGGCGAGCAACCGTGAA 1229
Db 280 aGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrVal 300

QY 1230 GCTCTATACATCAGTGGAGAAATGCTCAGCAGGATAGGGTGTGCTATCCCCACGATAT 1289
Db 300 sLeuTyrThrSerValGluAsnAlaGlnAspGlyValAlaIleProHisAspIle 320

QY 1290 CGATCTTGGTGATTGCGTGTGTCATTTCAGGATTATGACAACCAAGCATGAGCAGTGC 1349
Db 320 eAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspAr 340

QY 1350 GCCACCCCGCTCGCTCGGCCATCTCGGCTTTTCTGTTCTCGAGCAAAATGATGACT 1409
Db 340 gProThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLe 360

QY 1410 TTGCTCTCCCTCACTCAGCGAGTATGACCTCCACTTACGGGTGCTCACTGCGCC 1469
Db 360 uTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPyr 380

QY 1470 GGTATTATCTCGACAGCGTGAATTTGGTGAATTTGCGACTGGCGCGCAGGCGCTAGC 1529
Db 380 oValTyrIleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAl 400

QY 1530 CCATCGCTGTACTGTCCAAAGTCACCTCGAGGGCGGCGCCCTCCGACTGTTGAGCA 1589
Db 400 aArgSerLeuAspTrpSerLysValThrLeuAspGlyArgProLeuProThrValGluGl 420

QY 1590 ATATTCCAGACATCTTGTGTCTCCCTTCCGTCGTCGCAAGCTCTCTTTTGGGAGGCGCG 1649
Db 420 nTyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheThrGluAlaGl 440

QY 1650 CACAACAAAAGAGGTATCTTTATATATAATCTACTGTAGTGACCAAGATTTCTGAT 1709
Db 440 yThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeuIle 460

QY 1710 TGAATAATGCTCGCGCCATCGGTCGCCATTTCACTATACCAAGGCTTGGGCGCG 1769
Db 460 eGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrArgLeuGlyAlaGl 480

QY 1770 TCCGGTCCGCAATTTCTGCGCGCGGTTTGGCTCCACGCTCCGCGCTGCTGCTGGA 1829
Db 480 yProValAlaIleSerAlaAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuGlu 500

QY 1830 GGATCTTTTGTATATCCGGGCGGCGCACATTTGATGACTTCTGCGCTGAATCCG 1889
Db 500 uAspThrPheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysAr 520

QY 1890 CGCTTAGGCTCCAG 1905
Db 520 gAlaLeuGlyLeuGln 525

AA96094
ID AAR96094 standard; protein; 327 AA.
AC AAR96094;
XX
DT 16-OCT-2003 (revised)
DT 06-AUG-1996 (first entry)
XX
DE Hepatitis E virus (Mexico strain) SG3 antigen.
XX
KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW diagnosis; antigen.
XX
OS Hepatitis E virus; Mexico strain.
XX
FN WO9612807-A2.
XX
PD 02-MAY-1996.
XX
PF 23-OCT-1995; 95WO-US013703.
XX
PR 24-OCT-1994; 94US-00327952.
PR 13-OCT-1995; 95US-00542634.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Fuerst TR, Mcatee CP, Yarbough PO, Zhang Y;
XX
DR WPI; 1996-230608/23.
DR N-PSDB; AAT27112.
XX
PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
PT diagnostic reagents for determining HEV infection and in vaccines.
XX
PS Disclosure; Page 87-88; 125pp; English.
XX
CC SG3 antigen (AAR96094) comprises the C-terminal 327 amino acids of
CC hepatitis E virus (HEV) Mexico strain capsid protein (see also AAR96090).
CC It is obtd. inserting encoding DNA (AAT27112) into pBluescript and
CC ligating the cloned insert into vector pGEX for prodn. of SG3 in E. coli
CC cells. SG3 shows low antigenicity compared with capsid 62K antigen (see
CC also AAR96091-92 and AAR96101-04) from HEV Burma and Mexico strains that
CC are produced using a baculovirus expression system. (Updated on 16-OCT-
CC 2003 to standardise OS field)
XX
SQ Sequence 327 AA;
Alignment Scores:
Pred. No.: 9,85e-308 Length: 327
Score: 327.00 Matches: 327
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.62% Indels: 0
DB: 2 Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR96094 (1-327)
QY 997 GGGGCGGAGGAGTGGCGAGTGACCAACTGCAGCCACCCAGGTTTCATGAAGATCTC 1056
Db 1 GlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLysAspLeu 20
QY 1057 CACTTTACCGGCTTAATGGGCTAGGTGAAGTCGGCGGGGATAGCTCTAACATTACTT 1116
Db 21 HisPheThrGlyLeuAsnGlyValGlyGluValGlyArgGlyLeuAlaLeuThrLeuLeu 40
QY 1117 AACCTTGCTGACAGCTCTCGGCGGGTCCCGACAGAAATTAATTCGTGGCTGGCGGG 1176
Db 41 AsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuLeuSerSerAlaGlyGly 60
QY 1177 CAACCTGTTTTATTCCCGCGCGTGTCTCTCAGCAATGGCGAGCCCAACCGTGAAGTCTAT 1236
Db 61 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 80

QY 1237 ACATCAGTGGAGATGCTCAGCAGGATAAGGGTGTGTGCTATCCCCACGATATCGATCTT 1296
Db 81 ThrSerValGluAsnAlaGlnAspLysGlyValAlaAlaIleProHisAspLeuLeu 100
QY 1297 GGTGATTCGGTGTGTCATTGAGATTAAGCAACCCAGCATGACAGGATCGCCACC 1356
Db 101 GlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 120
QY 1357 CGGTGCGCTGGCCATCTCGGCTTTCTCTCCGAGCAAAATGATGTACTTTGGCTG 1416
Db 121 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuThrLeu 140
QY 1417 TCCCTCACTGCAGCCGAGTATGACCACTTACGGGTCTGCTAACCTGGCCGGTTTAT 1476
Db 141 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 160
QY 1477 ATCTCGACAGGTGACTTTGTGTAATGTTGGACTGGCGCGCAGCGGTAGCCGATCG 1536
Db 161 IleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValaAlaArgSer 180
QY 1537 CTTGACTGCTCCAAAGTCACCTCGACGGGGCCCTCCCGACTGTTGAGCAATATTC 1596
Db 181 LeuAspTyrSerLysValThrLeuAspGlyArgProLeuProThrValGluGlnTyrSer 200
QY 1597 AAGACATTTCTTTGTGCTCCCTTCGTGGCAAGCTCTCTTTTGGAGCGCCGCAACA 1656
Db 201 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 220
QY 1657 AAAGCAGGTTATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1716
Db 221 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeuIleGluAsn 240
QY 1717 GCTCGCGCCATCGGGTCCGCTTCAACCTATACACAGGCTTGGGGCGGTCGGGTC 1776
Db 241 AlaAlaGlyHisArgValAlaIleSerThrThrThrArgGluGlyAlaGlyProVal 260
QY 1777 GCCATTTCTGGCGCGCGGTTTGGCTCCACGCTCCGCTGGCTGCTGCTGGAGGATACT 1836
Db 261 AlaIleSerAlaAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuLeuGluAspThr 280
QY 1837 TTTGATTAATCGGGCGGGCGGCGACACATTTGATGACTTCTGCTGATGCGCGCTTTA 1896
Db 281 PheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysArgAlaLeu 300
QY 1897 GGCCTCCAGGTTGTGCTTCCAGTCAACTGTCGCTGAGCTCCAGCGCTTAAAGTTAAG 1956
Db 301 GlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysValLys 320
QY 1957 GTGGTAAACTCGGAGTTG 1977
Db 321 ValGlyLysThrArgGluLeu 327
RESULT 10
AAW35821
ID AAW35821 standard; protein; 327 AA.
XX
AC AAW35821;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 26-FEB-1998 (first entry)
XX
DE Hepatitis E virus Mexico strain peptide antigen SG3.
XX
KW Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody;
KW diagnosis; HEV.
XX
OS Hepatitis E virus; - Mexico strain.
XX
PN US5686239-A.
XX
PD 11-NOV-1997.
XX


```
PF 09-MAY-1994; 94US-00240049.
XX
PR 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-0036672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 05-APR-1991; 91US-00681078.
PR 17-JAN-1992; 92US-00822335.
PR 20-APR-1992; 92US-00870985.
PR 01-MAY-1992; 92US-00876941.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
XX Tam AW, Reyes GR, Yarbough PO;
XX WPI; 1997-558132/51.
DR N-PSDB; AAT96966.
XX
PT Diagnosis of hepatitis E virus Burma and Mexico strain infection - by
PT immunoassay with hepatitis E virus peptide antigens.
XX
XX Claim 1; Col 37-40; 36pp; English.
XX
XX A method has been developed for detecting hepatitis E virus (HEV)
CC antibodies (Ab). The method comprises: (a) reacting a serum sample with a
CC HEV peptide antigen; and (b) examining the peptide for the presence of
CC bound Ab, where the presence of bound Ab indicates the presence of HEV
CC Ab. The present sequence represents a specifically claimed HEV peptide
CC antigen. The method can be used to diagnose infection with the
CC enterically transmitted non-A/non-B viral hepatitis agent HEV,
CC specifically the HEV Burma and Mexico strains. (Updated on 25-MAR-2003 to
CC correct PF field.) (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 327 AA;
XX
Alignment Scores:
Pred. No.: 9,85e-308 Length: 327
Score: 327.00 Matches: 327
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.62% Indels: 0
DB: 2 Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAW35821 (1-327)
QY 997 GGGCCGACGGAGCTCGGAGCTGACCAACTGACGACCCAGCGTTCATGAAGAATCTC 1056
Db 1 GlyAlaAspGlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAspLeu 20
QY 1057 CACTTTACCGCCTTAATGGGTAGGTGAGTCGGCGCGGATAGCTTAACATTTACTT 1116
Db 21 HisPheThrGlyLeuAsnGlyValGlyGluValGlyArgGlyIleAlaLeuThrLeuLeu 40
QY 1117 AACCTTCTGACAGCTCCTCGCGGCTCCCGCAGAAATTAATTCGTGCGCTGGCGGG 1176
Db 41 AsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGly 60
QY 1177 CAATGTTTATTCGCCCGCGGTGTCTCAGCCAATGCGAGCAACCGTGAAGCTCTAT 1236
Db 61 GlnLeuPheThrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuThr 80
QY 1237 ACATCAGTGGAGATGCTCAGCAGATAAGGTTCTGCTATCCCGCAGGATATGATCTT 1296
Db 81 ThrSerValGluAsnAlaGlnAspLysGlyValAlaIleProHisAspIleAspLeu 100
QY 1297 GGTGATTGCGGTGGTCATTGAGATTATGACAAACAGCATGAGCAGGATCGGCCAC 1356
Db 101 GlyAspSerArgValValIleGlnAspTyrrAspAsnGlnHisGluGlnAspArgProThr 120
QY 1357 CCGTCGCGCGCATCTCGGCCCTTTTCTGCTTCTCCGAGCAAAATGATGACTTTGGCTG 1416
Db 121 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 140
```

```
QY 1417 TCCTCACTGCAGCCGAGTATGACCACTTCACCTTACGGGTGCTCAACTGGCCCGGTTTAT 1476
Db 141 SerLeuThrAlaAlaGluTyrrAspGlnSerThrTyrrGlySerSerThrGlyProValTyrr 160
QY 1477 ATCTCGGACAGCGTGAATTTGGTGAATGTTGCGACTGCGCGCAGGCGCGTAGCCCGATCG 1536
Db 161 IleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 180
QY 1537 CTGACTGTCGTCMAAGTACCTCGACGGGGGGCCCTCCCGACTGTTGAGCAATATTCC 1596
Db 181 LeuAspTrpSerLysValThrLeuAspGlyArgProLeuProThrValGluGlnTyrrSer 200
QY 1597 AAGACATTCTTGCTCCCTCCCTCGTGGCAAGCTCTCTTTGGGAGGCGCGCACAAACA 1656
Db 201 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 220
QY 1657 AAGCAGGTATCTTATAATTAATACTACTACTAGTACCAAGATTCTGATTGAAAAAT 1716
Db 221 LysAlaGlyTyrrProTyrrAsnTyrrAsnThrThrAlaSerAspGlnIleLeuIleGluAsn 240
QY 1717 GCTGCCGCGCATCGGTGCGCATTTCAACCTATACCAACCGAGCTTGGGCGCGGTCCGGTC 1776
Db 241 AlaAlaGlyHisArgValAlaIleSerThrTyrrThrArgLeuGlyAlaGlyProVal 260
QY 1777 GCCATTTCTCGCGCGCGGTTTGGCTCCACGCTCCGCGCTCGCTGCTGGAGGATACT 1836
Db 261 AlaIleSerAlaAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuLeuGluAspThr 280
QY 1837 TTTGATTATCCGGGCGCGGCGCACACATTTGATGACTTCTGCCCTGAATGCCGCGCTTTA 1896
Db 281 PheAspTyrrProGlyArgAlaHisThrPheAspAspPheCysProGluCysArgAlaLeu 300
QY 1897 GGCTTCAGGGTGTGCTTTCCAGTCAACTGTGCTGAGCTCCAGCGCTTAAAGTTAAG 1956
Db 301 GlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysValLys 320
QY 1957 GTGGGTAAAACTCGGAGTTG 1977
Db 321 ValGlyLysThrArgGluLeu 327
RESULT 11
ABJ04808
ID ABJ04808 standard; protein; 659 AA.
AC ABJ04808;
XX
DT 07-NOV-2002 (first entry)
XX
DE Hepatitis E virus Mexico strain ORF2 capsid gene.
XX
KW Avian hepatitis E virus; vaccine; hepatitis E; ORF2; zoonosis; virucide;
KW avian hepatitis-splenomegaly syndrome; HEV; HS syndrome.
XX
OS Hepatitis E virus.
XX
PN W0200253712-A2.
XX
PD 11-JUL-2002.
XX
PF 04-JAN-2002; 2002WO-US0000215.
XX
PR 05-JAN-2001; 2001US-0259846P.
PR 31-DEC-2001; 2001US-00029840.
XX
PA (VIRG ) VIRGINIA TECH INTELLECTUAL PROPERTIES.
XX
PI Meng X, Haqshenas G, Huang F;
XX
XX WPI; 2002-548085/58.
XX
XX Novel isolated avian hepatitis E virus useful in a vaccine for protecting
PT an avian or mammalian species from viral infection or hepatitis-
```


QY 490 CTGCGCTGACGAGCGGTACTATACATCATTTATGGCCACAGAGCGCTCCAAATTATGCA 549
Db |||||||
53 LeuProLeuGlnAspGlyThrAsnThrHisileMetAlaThrGluAlaSerAsnTyrAla 72
QY 550 CAGTACCGGGTTGCCGCGCTACTATCCGTTACCGGCCCTAGTGCCTTAATGCGAGTTGGA 609
Db |||||||
73 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 92
QY 610 GGCTATGCTATATCATTTCTTCTGGCTCAAAACCAACCAACCCCTACATCTCTGTTGAC 669
Db |||||||
93 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 112
QY 670 ATGAATTCCATTTACTTCCACTGATGTCAGGATTTCTTGTCAACCTGGCATAGCATCTGAA 729
Db |||||||
113 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 132
QY 730 TTGCTCATCCCAAGGAGCGCTTCACTACCGCATCAAGTTGGCGCTCGGTTGAGACA 789
Db |||||||
133 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 152
QY 790 TCTGCTGTGTGCTGAGGAGGAGCCACCTCCGGTCTTGTGTCATGTTATGCATATGCTCT 849
Db |||||||
153 SerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 172
QY 850 CCAGTTAACTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGGACTTTGCC 909
Db |||||||
173 ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla 192
QY 910 TTAGAGCTTGAGTTTCGAATCTCACC 936
Db |||||||
193 LeuGluLeuGluPheArgAsnLeuThr 201
RESULT 13
AAR70323
ID AAR70323 standard; protein; 660 AA.
XX AC AAR70323;
XX XX
DT 25-MAR-2003 (revised)
DT 28-SEP-1995 (first entry)
XX XX
DE Hepatitis E virus ORF2 protein.
XX XX
KW HEV; ORF2; antigen; vaccine; immunogen.
XX XX
OS Hepatitis E virus.
XX XX
PN W09508632-A1.
XX XX
PD 30-MAR-1995.
XX XX
PF 23-SEP-1994; 94WO-AU000572.
XX XX
PR 24-SEP-1993; 93AU-00001423.
PR 15-DEC-1993; 93AU-00002964.
XX XX
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX XX
PI Anderson DA, Locarnini SA, Torresi J, Li F, Hui Z;
XX XX
DR WPI; 1995-139601/18.
DR N-PSDE; AAQ86592.
XX XX
PT Antigens of hepatitis E virus (HEV) - selectively immuno-reactive to
XX convalescent and/or acute phase circulating antibodies to HEV.
XX PS Disclosure; Page 39-41; 78pp; English.
XX XX
CC RNA from an HEV strain isolated in the Xinjiang region of China was
CC subjected to RT-PCR to obtain fragments corresp. to ORF2 (given in
CC AAQ86592) and ORF3 (AAQ86593) that encoded antigenic proteins (AAR70323-
CC 24, respectively). DNA fragments were manipulated into pGEX vectors for
CC production of GST fusion proteins in E. coli. Applications include HEV

CC infection diagnosis, therapy and vaccine development. (Updated on 25-MAR-
CC 2003 to correct PN field.)
SQ Sequence 660 AA;
Alignment Scores:
Pred. No.: 4,23e-135 Length: 660
Score: 149.00 Matches: 149
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.61% Indels: 0
DB: Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR70323 (1-660)
QY 490 CTGCGCTGACGAGCGGTACTATACATCATTTATGGCCACAGAGCGCTCCAAATTATGCA 549
Db |||||||
164 LeuProLeuGlnAspGlyThrAsnThrHisileMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGGTTGCCGCGCTACTATCCGTTACCGGCCCTAGTGCCTTAATGCGAGTTGGA 609
Db |||||||
184 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203
QY 610 GGCTATGCTATATCATTTCTTCTGGCTCAAAACCAACCAACCCCTACATCTCTGTTGAC 669
Db |||||||
204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCCATTTACTTCCACTGATGTCAGGATTTCTTGTCAACCTGGCATAGCATCTGAA 729
Db |||||||
224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 TTGCTCATCCCAAGGAGCGCTTCACTACCGCATCAAGTTGGCGCTCGGTTGAGACA 789
Db |||||||
244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 263
QY 790 TCTGCTGTGTGCTGAGGAGGAGCCACCTCCGGTCTTGTGTCATGTTATGCATATGCTCTCT 849
Db |||||||
264 SerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283
QY 850 CCAGTTAACTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGGACTTTGCC 909
Db |||||||
284 ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla 303
QY 910 TTAGAGCTTGAGTTTCGAATCTCACC 936
Db |||||||
304 LeuGluLeuGluPheArgAsnLeuThr 312
RESULT 14
AAR91814
ID AAR91814 standard; protein; 660 AA.
XX AC AAR91814;
XX XX
DT 26-NOV-1996 (first entry)
XX XX
DE Hepatitis E virus strain SAR-55 ORF-2.
XX XX
KW Hepatitis E virus; HEV; SAR-55 strain; enteric transmission;
KW structural region; antigen; detection; antibody; vaccine; immunisation;
XX XX
OS Hepatitis E virus.
XX XX
PN W09610590-A2.
XX XX
PD 11-APR-1996.
XX XX
PF 03-OCT-1995; 95WO-US013102.
XX XX
PR 03-OCT-1994; 94US-00316765.
XX XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX XX

```
PI Tsarev SA, Emerson SU, Purcell RH;
XX WPI; 1996-209320/21.
DR N-PSDB; AAT27394.
XX
PT Isolated and purified hepatitis E virus strain SAR-55 DNA - encodes
PT antigenic protein useful in diagnosis, prophylaxis and treatment of
PT hepatitis E virus infection.
XX
XX Claim 6; Page 13-15; 121pp; English.
XX
XX The present sequence is the protein prod. of ORF-2 from the hepatitis E
XX virus (HEV) strain SAR-55, which was implicated in an enterically
XX transmitted non-A, non-B hepatitis in Pakistan. The protein encoded by
XX the structural region of the virus (i.e. ORF-2), which is capable of
XX forming HEV like particles, is useful for the detection of HEV antibodies
XX (pref. IgG or IgM) in blood, plasma, sera, cerebrospinal fluid, tissue,
XX urine or pleural fluid. The protein, and anti-HEV antibodies generated
XX using the protein, can also be used in vaccines for immunising an animal
XX against HEV infection. The protein is identified as a band of greater
XX than 50 kD following SDS-PAGE of cell lysates of insect cells infected
XX with a HEV ORF-2 contg. baculovirus, i.e. the claimed recombinant
XX expression vectors pPIC9-1779, -1780 and -1781
XX
XX Sequence 660 AA;
XX
XX Alignment Scores:
XX Pred. No.: 4,238-135 Length: 660
XX Score: 149.00 Matches: 149
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 22.61% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAR91814 (1-660)
XX
QY 490 CTGCGCGTGCAGAGCGGTACTAATCTACATTATGGCCACAGAGCGCTCCAATTATGCA 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGGTGGCCGGCGCTACTATCCGTTACCGGCCCTAGTGGCTAATCGAGTTGA 609
Db 184 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203
QY 610 GCCTATGCTATATCCATTCTTTCTGGCCTCAAAACACCAACCCCTACATCTGTGAC 669
Db 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCCATTACTTCCACTGATGTCAGGATTCTTGTTCACCTGGCATAGCATCTGAA 729
Db 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 TTGGTTCATCCCAAGGAGCGCTTCACTACCGCAATCAAGTTGGCGCTCGGTGAGACA 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 263
QY 790 TCTGTTGTTGTCGAGGAGGAGCGACCTCCGGTCTTGTTCATGTTATGCATATGGCTCT 849
Db 264 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283
QY 850 CCAGTTAACTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGCACTTGGC 909
Db 284 ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla 303
QY 910 TTAGAGCTTGAGTTTCGGCAATCTCACC 936
Db 304 LeuGluLeuGluPheArgAsnLeuThr 312
XX
XX RESULT 15
XX AAW76369
XX ID AAW76369 standard; protein; 660 AA.
XX AC AAW76369;
```

```
XX 03-DEC-1998 (first entry)
XX Hepatitis E virus hollow particle protein #2.
XX Hollow particle protein; virus; antibody; detection; immunoassay;
XX infection.
XX Hepatitis virus.
XX JP10234383-A.
XX 08-SEP-1998.
XX 28-FEB-1997; 97JP-00062445.
XX 28-FEB-1997; 97JP-00062445.
XX (DENK-) DENKA SEIKEN KK.
XX (KOKU-) KOKURITSU YODO BISEI KENKYUSHO.
XX WPI; 1998-535037/46.
XX N-PSDB; AAV61888.
XX Hepatitis E virus hollow particle poly:peptide(s) and nucleic acids
XX encoding it - useful for more accurate detection of HEV in samples, using
XX immuno-assays and nucleic acid hybridisation.
XX Claim 13; Page 24-26; 29pp; Japanese.
XX This sequence represents a Hepatitis E viral hollow particle protein.
XX This polypeptide can be used to raise antibodies to detect HEV infection
XX in samples, e.g. by immuno-assay based techniques, and the nucleic acid
XX can be used for the same in nucleic acid hybridisation assays. The
XX polypeptides and nucleic acids allow more accurate detection of HEV than
XX previously possible
XX Sequence 660 AA;
XX
XX Alignment Scores:
XX Pred. No.: 4,238-135 Length: 660
XX Score: 149.00 Matches: 149
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 22.61% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-09-851-410A-10_COPY_5117_7096 (1-1980) x AAW76369 (1-660)
XX
QY 490 CTGCGCGTGCAGAGCGGTACTAATCTACATTATGGCCACAGAGCGCTCCAATTATGCA 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGGTGGCCGGCGCTACTATCCGTTACCGGCCCTAGTGGCTAATCGAGTTGA 609
Db 184 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203
QY 610 GCCTATGCTATATCCATTCTTTCTGGCCTCAAAACACCAACCCCTACATCTGTGAC 669
Db 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCCATTACTTCCACTGATGTCAGGATTCTTGTTCACCTGGCATAGCATCTGAA 729
Db 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 TTGGTTCATCCCAAGGAGCGCTTCACTACCGCAATCAAGTTGGCGCTCGGTGAGACA 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 263
QY 790 TCTGTTGTTGTCGAGGAGGAGCGACCTCCGGTCTTGTTCATGTTATGCATATGGCTCT 849
Db 264 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283
QY 850 CCAGTTAACTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGCACTTGGC 909
Db 284 ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla 303
QY 910 TTAGAGCTTGAGTTTCGGCAATCTCACC 936
Db 304 LeuGluLeuGluPheArgAsnLeuThr 312
XX
XX RESULT 15
XX AAW76369
XX ID AAW76369 standard; protein; 660 AA.
XX AC AAW76369;
```

Qy	850	CCAGTTAACTCCTATACCAATACCCCTTATACCGGTGCCCTTGCTTACTGGACTTTGCC	909
Db	284	ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaIeuGlyLeuLeuAspPheAla	303
Qy	910	TTAGAGCTTGAGTTTCGCAATCTCACC	936
Db	304	LeuGluLeuGluPheArgAsnLeuThr	312

Search completed: August 16, 2004, 13:48:11
Job time : 114.682 secs

Blank sheet

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.21% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VHHWH2 (1-660)

```
QY 490 CTGCGCGTGCAGGACGGTACTTAATCTCACAATTATGGCCACAGAGCGCTCCAAATTATGCA 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGGTGGCCGGCTACTATCCGTTACCGGCCCTAGTCCCTAATGCAAGTTGGA 609
Db 184 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203
QY 610 GGCTATGCTATATCATTTCTTCTGGCTCAAAACACCAACCCCTACATCTGTTGAC 669
Db 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCCTATTTACTTCCACTGTGTCAGGATTCTTGTTCAACTGGCATAGCATCTGAA 729
Db 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 TTGCTATCCCAAGCAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCGGTTGACACA 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThr 263
QY 790 TCTGGTCTGCTGAGGAGGAGCCACCTCCGGTCTTGTCAATGTTATGCATACATGGCTCT 849
Db 264 SerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283
```

RESULT 3

C44212
structural protein 1 - hepatitis E virus (strain Mexico)
C;Species: hepatitis E virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C;Accession: C44212
R;Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.
Virology 191, 550-558, 1992
A;Title: Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HE)
A;Reference number: A44212; MUID:93079857; PMID:1448913
A;Accession: C44212
A;Molecule type: genomic RNA
A;Residues: 1-123 <HUA>
A;Cross-references: GB:M74506; NID:g330017; PIDN:AAA45731.1; PID:g330019
C;Superfamily: hepatitis E virus structural protein 1
C;Keywords: structural protein

Alignment Scores:
Pred. No.: 5.5e-101 Length: 123
Score: 109.00 Matches: 109
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.54% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x C44212 (1-123)

```
QY 2 TGGCCCTAGCCCTCTTTGCTGTGTTCTCTTGTGTTCTGCCTATGTTGCCCGCGGCAC 61
Db 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGACCGGTGACCGCTCTGGCCGCGCTGTCGGCGGCGCAGCGGGTACCGCGGTGGTT 121
Db 35 ArgProValSerArgLeuAlaValAlaValGlyGlyAlaAlaValProAlaValVal 54
QY 122 TCTGGGTGACCGGTGATTCTCAGCCCTTCGCAATCCCCCTATATCATCAACCAACC 181
Db 55 SerGlyValThrGlyLeuLeuLeuSerProSerGlnSerProIlePheIleGlnProThr 74
QY 182 CTTTGGCCCGACAGTTGCGCGCTCGCTCGGGTCTGACCTCGCTCGCCAAACAGCC 241
Db 75 ProLeuProGlnThrLeuProLeuArgProGlyLeuAspLeuAlaPheAlaAsnGlnPro 94
```

```
QY 242 GGCCACTTGGCTCCACTTGGCGAGATCAGGCCACAGCCCGCTCGCTCCCTCCGTCGCC 301
Db 95 GlyHisLeuAlaProLeuGlyGluIleArgProSerAlaProLeuProProValAla 114
QY 302 GACCTGCCACACCGGGGCTGCGGCGC 328
Db 115 AspLeuProGlnProGlyLeuArgArg 123
```

RESULT 4

VHHWH2
structural protein 1 - hepatitis E virus (strain Burma)
C;Species: hepatitis E virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 23-Jul-1999
C;Accession: B40778; A40236
R;Tam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.C.; Bradley, D.W.; Fry, K.E.; Reyes, G.F.
Virology 185, 120-131, 1991
A;Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length vir
A;Reference number: A40778; MUID:92024067; PMID:1926770
A;Accession: B40778
A;Molecule type: genomic RNA
A;Residues: 1-123 <TAM>
A;Cross-references: GB:M73218; NID:g330023; PIDN:AAA45735.1; PID:g330025
R;Ray, R.; Jameel, S.; Manivel, V.; Ray, R.
Virology 189, 359-362, 1992
A;Title: Indian hepatitis E virus shows a major deletion in the small open reading frame.
A;Reference number: A40236; MUID:92295577; PMID:1534953
A;Accession: A40236
A;Molecule type: genomic RNA
A;Residues: 1-57, 'p', 59-102, 'p', 104-123 <RAY>
C;Superfamily: hepatitis E virus structural protein 1
C;Keywords: structural protein

Alignment Scores:
Pred. No.: 1.9e-52 Length: 123
Score: 61.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.26% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VHHWH2 (1-123)

```
QY 2 TGGCCCTAGCCCTCTTTGCTGTGTTCTCTTGTGTTCTGCCTATGTTGCCCGCGGCAC 61
Db 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGACCGGTGACCGCTCTGGCCGCGCTGTCGGCGGCGCAGCGGGTACCGCGGTGGTT 121
Db 35 ArgProValSerArgLeuAlaValAlaValGlyGlyAlaAlaValProAlaValVal 54
QY 122 TCTGGGTGACCGGTGATTCTCAGCCCTTCGCAATCCCCCTATATCATCAACCAACC 181
Db 55 SerGlyValThrGlyLeuIleLeuSerProSerGlnSerProIlePheIleGlnProThr 74
QY 182 CCT 184
Db 75 Pro 75
```

RESULT 5

B40236
structural protein 1 - hepatitis E virus (strain Indian)
C;Species: hepatitis E virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: B40236
R;Ray, R.; Jameel, S.; Manivel, V.; Ray, R.
Virology 189, 359-362, 1992
A;Title: Indian hepatitis E virus shows a major deletion in the small open reading frame.
A;Reference number: A40236; MUID:92295577; PMID:1534953
A;Accession: B40236
A;Molecule type: genomic RNA
A;Residues: 1-41 <RAY>
C;Superfamily: hepatitis E virus structural protein 1

C;Keywords: structural protein

Alignment Scores:
Pred. No.: 0.000778 Length: 41
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.97% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x B40236 (1-41)

QY 2 TCGGCGCTAGGCGCTCTTTGCGTGTGCTCTCTGTTTC 40

Db 15 CysAlaLeuGlyLeuPheCysSerCysPhe 27

RESULT 6

T36459

hypothetical protein SCF43A.36 - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T36459

R;Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1999

A;Reference number: Z21598

A;Accession: T36459

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-266 <SEE>

A;Cross-references: EMBL:AL096837; PIDN:CAB48923.1; GSPDB:GN00070; SCOEDB:SCF43A.36

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SCF43A.36

Alignment Scores:
Pred. No.: 6.51 Length: 266
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x T36459 (1-266)

QY 76 AGCGCTGACCGCTCGGTGCGGGGCA 50

Db 100 ThrAlaAspArgSerValAlaArgAla 108

RESULT 7

G69580

acetyl-CoA carboxylase (EC 6.4.1.2), carboxyltransferase alpha chain - Bacillus subtilis

C;Species: Bacillus subtilis

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C;Accession: G69580

R;Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berthel

A.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Konningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols,

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel

Y, M.; Ogawa, K.; Ogiwara, K.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: G69580

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-325 <KUN>

A;Cross-references: GB:Z99118; GB:AL009126; NID:g2635200; PIDN:CAB14880.1; PID:g2635385

A;Experimental source: strain 168

C;Genetics:

A;Gene: accA

C;Function:

A;Pathway: fatty acid biosynthesis

C;Superfamily: acetyl-CoA carboxylase, carboxyltransferase alpha chain

C;Keywords: fatty acid biosynthesis, ligase

Alignment Scores:
Pred. No.: 6.31 Length: 325
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x G69580 (1-325)

QY 294 GGAGGACGGAGGGGCGCTGGCGCTG 268

Db 199 GlyGlySerGlyGlyAlaLeuGlyLeu 207

RESULT 8

H58208

protamine II - black rat snake (fragment)

C;Species: Elaphe obsoleta quadrivittata

C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 07-May-1999

C;Accession: H58208

R;Hunt, J.G.; Kasinsky, H.E.; Eisey, R.M.; Wright, C.L.; Rice, P.; Bell, J.B.; Sharp, D.

J. Biol. Chem. 271, 23547-23557, 1996

A;Title: Protamines of reptiles.

A;Reference number: A58208; MUID:963994458; PMID:8798564

A;Accession: H58208

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-41 <HUN>

C;Superfamily: sperm histone

Alignment Scores:
Pred. No.: 88.5 Length: 41
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x H58208 (1-41)

QY 76 TCTGGCGCGCTCGTGGGCGGCGC 99

Db 13 SerGlyArgArgGlyArgArg 20

RESULT 9

S19708

hypothetical protein C - Herpetosiphon aurantiacus

C;Species: Herpetosiphon aurantiacus

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Jan-1995

C;Accession: S19708

R;Brdmann, D.; Duesterhoeft, A.; Kroeger, M.

Eur. J. Biochem. 202, 1247-1256, 1991

A;Title: Cloning and molecular characterization of the HgiCI restriction/ modification

A;Reference number: S19706; MUID:92111503; PMID:1662609

A;Accession: S19708

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-74 <KRO>

A;Cross-references: EMBL:X55138

Alignment Scores:
Pred. No.: 81 Length: 74
Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x S19708 (1-74)

Qy 827 ACAAGACGGAGGTGGCTTCCTCC 804
|||||
Db 64 ThrArgProGluValAlaSer 71
|||||

RESULT 10
C61547
hypothetical protein 1 (px region) - human T-cell lymphotropic virus type 1 (isolate HAM
C;Species: human T-cell lymphotropic virus type 1, HTLV-1
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 08-Oct-1999
C;Accession: C61547
R;Tsujimoto, A.; Teruuchi, T.; Imamura, J.; Shimotohno, K.; Miyoshi, I.; Miwa, M.
Mol. Biol. Med. 5, 29-42, 1988
A;Title: Nucleotide sequence analysis of a provirus derived from HTLV-1-associated myelo
A;Reference number: A61547; MUID:88232270; PMID:2897612
A;Accession: C61547
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-99 <TSU>
A;Cross-references: GB:M37301; NID:G541634; PIDN:AAA45390.1; PID:G541637

Alignment Scores:
Pred. No.: 77.5 Length: 99
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x C61547 (1-99)

Qy 22 CTGTGTCTCTCTCTGTCTCTGCCT 45
|||||
Db 52 LeuLeuPheLeuLeuPheLeuPro 59
|||||

RESULT 11
C90345
hypothetical protein SS01819 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: C90345
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayes, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Zensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: C90345
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-104 <KUR>
A;Cross-references: GB:AE006641; NID:gl3815078; PIDN:AAK42018.1; GSPDB:GN00155
C;Genetics:
A;Gene: SS01819

Alignment Scores:
Pred. No.: 76.9 Length: 104
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x C90345 (1-104)

Qy 1115 TTAACCTGTGCACGCTCTCG 1138
|||||

Db 16 LeuThrLeuLeuThrArgSerSer 23

RESULT 12
D75495
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C;Accession: D75495
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Mal
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: D75495
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-142 <WHI>
A;Cross-references: GB:AE001921; GB:AE000513; NID:G6458330; PIDN:AAF10219.1; PID:G6458341
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0638
A;Map position: 1
C;Superfamily: Deinococcus radiodurans hypothetical protein DR0638

Alignment Scores:
Pred. No.: 73.4 Length: 142
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x D75495 (1-142)

Qy 305 GGTCGGGACGGAGCGAGCGAG 282
|||||
Db 97 GlyArgArgGluAlaAlaGlu 104
|||||

RESULT 13
JQ1865
hypothetical 15.1K protein - bovine adenovirus 3
C;Species: Mastadenovirus bos3 (bovine adenovirus 3)
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C;Accession: JQ1865
R;Mittal, S.K.; Prevec, L.; Babiuk, L.A.; Graham, F.L.
J. Gen. Virol. 73, 3295-3300, 1992
A;Title: Sequence analysis of bovine adenovirus type 3 early region 3 and fibre protein 3
A;Reference number: PQ0499; MUID:93107871; PMID:1459367
A;Accession: JQ1865
A;Molecule type: DNA
A;Residues: 1-142 <MIT>
A;Cross-references: GB:M12928
A;Experimental source: strain WBR-1

Alignment Scores:
Pred. No.: 73.4 Length: 142
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x JQ1865 (1-142)

Qy 1022 CCACACTGCAGCCACCAGGTTC 1045
|||||
Db 57 ProGlnLeuGlnProGlySer 64
|||||

RESULT 14
A46181
pX-rer-orf I (alternatively spliced) - human T-cell lymphotropic virus type 1
C;Species: human T-cell lymphotropic virus type 1, HTLV-1

C;Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000
C;Accession: A46181; B46181
R;Korainik, I.J.; Gessain, A.; Klotman, M.E.; Lo Monaco, A.; Berneman, Z.N.; Franchini,
Proc. Natl. Acad. Sci. U.S.A. 89, 8813-8817, 1992
A;Title: Protein isoforms encoded by the pX region of human T-cell leukemia/lymphotropic
A;Reference number: A46181; MUID:92409607; PMID:1528897

A;Accession: A46181
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-152 <KOR>
A;Note: sequence extracted from NCBI backbone (NCBIP:114304)
A;Accession: B46181
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 54-152 <K02>
A;Note: sequence extracted from NCBI backbone (NCBIP:114306)

Alignment Scores:			
Pred. NO.:	72.7	Length:	152
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	2	Gaps:	0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x A46181 (1-152)

QY 22 CTGTTGTCCTCTGTTCTGCGCT 45
DB 105 LeuLeuPheLeuLeuPheLeuPro 112

RESULT 15

D87322
hypothetical protein CC0590 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: D87322
R;Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
R. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: D87322
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-153 <STO>
A;Cross-references: GB:AE005673; NID:gi3421790; PIDN:AAK22576.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC0590

Alignment Scores:			
Pred. NO.:	72.6	Length:	153
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	2	Gaps:	0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x D87322 (1-153)

QY 313 GCCGGGCTCGCGCGCTGACGGCT 336
DB 109 AlaGlyAlaAlaAlaLeuThra 116

Search completed: August 16, 2004, 13:57:47
Job time : 39.2263 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:36:13 ; Search time 18.2362 Seconds
(without alignments)
11307.070 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 659

Sequence: 1 ATGGCCCTAGGCTCTTTT.....GTAAACTCGGAGTGTAG 1980

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Word size: 1

Total number of hits satisfying chosen parameters: 283186

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09851410/runat_16082004_125716_9625/app.query.fasta_1.4238
-DB=SwissProt 42 -QFMT=fastan -SUFFIX=oli.rsp -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410@cgn 1 1.43 @runat_16082004_125716_9625 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	659	100.0	659	VST2_HEVME	Q03500 hepatitis e
2	149	22.6	660	VST2_HEVFA	P33426 hepatitis e
3	124	18.8	485	VST2_HEVRH	Q00270 hepatitis e
4	124	18.8	660	VST2_HEVMY	Q04611 hepatitis e
5	120	18.2	660	VST2_HEVBU	P29326 hepatitis e
6	109	16.5	123	VST1_HEVME	Q03499 hepatitis e
7	61	9.3	123	VST1_HEVBU	P29325 hepatitis e
8	59	9.0	123	VST1_HEVMY	Q04612 hepatitis e
9	9	1.4	325	ACCA_BACSU	Q34847 bacillus su
10	8	1.2	121	RL7_CLOPE	P08498 human adeno
11	8	1.2	131	E315_ADE05	P28220 sus scrofa
12	8	1.2	153	SORR_PIG	P28229 herpes slmp
13	8	1.2	233	UL03_HSV2H	P51249 porphyra pu
14	8	1.2	241	RR2_PORPU	P29786 aedes aegypt
15	8	1.2	254	TRY3_AEDAE	Q09952 caenorhabdi
16	8	1.2	267	YSR4_CAEEL	Q9fuz2 arabidopsis
17	8	1.2	273	DEFC_ARATH	P46316 anthranilic
18	8	1.2	322	ACCA_ANTSP	

c	19	8	1.2	324	1	ACCA_CYACA	O19903 cyanidium c
c	20	8	1.2	324	1	ACCA_PORPU	P51371 porphyra pu
c	21	8	1.2	326	1	ACCA_SYNY3	P74638 synechocyst
c	22	8	1.2	327	1	ACCA_SYNP7	Q54766 synechococc
c	23	8	1.2	331	1	MANI_MOUSE	Q9wu40 mus musculus
c	24	8	1.2	386	1	DGT1_ZYMMO	Q9gm70 zymomonas m
c	25	8	1.2	397	1	S17A_RABIT	Q9m700 oryctolagus
c	26	8	1.2	414	1	K193_MOUSE	Q9czc8 mus musculus
c	27	8	1.2	448	1	PDFT_CANAL	P78589 candida alb
c	28	8	1.2	449	1	CAPT_PSEPU	P28269 pseudomonas
c	29	8	1.2	459	1	P3A2_STRPU	Q04073 strongyloce
c	30	8	1.2	474	1	RCA_ARATH	P10896 arabidopsis
c	31	8	1.2	505	1	EX7L_CAUAR	Q9a649 caulobacter
c	32	8	1.2	507	1	MKR3_HUMAN	Q13064 homo sapien
c	33	8	1.2	572	1	PDFT_USTMA	Q92459 ustilago ma
c	34	8	1.2	584	1	ACES_RABIT	Q29499 oryctolagus
c	35	8	1.2	614	1	ACES_HUMAN	P22303 homo sapien
c	36	8	1.2	614	1	ACES_RAT	P37136 rattus norv
c	37	8	1.2	732	1	PSAB_CYAME	Q85fy6 cyanidiosch
c	38	8	1.2	735	1	DUR3_YEAST	P33413 saccharomyc
c	39	8	1.2	800	1	T2D4_HUMAN	Q15542 homo sapien
c	40	8	1.2	928	1	CHS2_EXODE	P30601 exophiala d
c	41	8	1.2	1040	1	YEGN_ECOLI	P76398 escherichia
c	42	8	1.2	1625	1	CTPI_MYCTU	Q10900 mycobacteri
c	43	8	1.2	1783	1	RAA3_CHLRE	Q9fec4 chlamydomon
c	44	8	1.2	3969	1	HRX_HUMAN	Q03164 homo sapien
c	45	7	1.1	26	1	MEL_APIDO	P01502 apis dorsat

ALIGNMENTS

RESULT 1

VST2_HEVME VST2_HEVME STANDARD; PRT; 659 AA.

AC Q03500;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor.
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31768;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93079857; PubMed=1448913;
RA Huang C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E.,
RA Bradley D.W., Tam A.W., Reyes G.R.;
RT "Molecular cloning and sequencing of the Mexico isolate of hepatitis
E virus (HEV).";
RL Virology 191:550-558(1992).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M74506; AAA45732.1; -
CC PIR; B44212; B44212.
CC InterPro; IPR004261; SP2.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF03014; SP2; 1.
CC Signal.
CC CHAIN 1 22 BY SIMILARITY.
FT SIGNAL 23 659 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 659 AA; 70640 MW; CF75E75EFD8FBE2C CRC64;

Alignment Scores:

Pred. No.:	0	Length:	659
Score:	659.00	Matches:	659
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-851-410A-10 COPY 5117 7096 (1-1980) x VST2 HEVME (1-659)

Qy	1	ATGCGCCTAGCCCTCTTTTGTGTTGTTCCTCTCTGTTTCTGCTATGTTGCCGCGCA	60
Db	1	MetArgProArgProLeuLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro	20
Qy	61	CCGACCGTCAAGCCTCTGGCCGCGTGTGTGGCGCGCAGCGGGGTACCGCGGTGT	120
Db	21	ProThrGlyGlnProSerGlyArgArgGlyArgArgSerGlyGlyThrGlyGlyGly	40
Qy	121	TTCTGGGGTGACCGGGTTGATTCTCAGCGCCTTCGCAATCCCTATATTTCATCCAA	180
Db	41	PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn	60
Qy	181	CCCTTTGCCCCAGAGCTTGCCTGCTCGGTCTGGACCTCGCTTCGCCAAACGAGCC	240
Db	61	ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla	80
Qy	241	CGGCACTTGGCTCCACATTGGCGAGATCAGGCCAGCGCCCTCGCTCCGCTCCCGTCG	300
Db	81	ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg	100
Qy	301	CGACCTGCCACAGCCGGGGCTGCGCGCTGACGGCTGTGGCGCCTGCCCATCACACCTCA	360
Db	101	ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer	120
Qy	361	CCCGTCCCGAGCTTGATTCTCGCGGTGAATCTACGCCCGCAGATATAATTTGCTACT	420
Db	121	ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr	140
Qy	421	TCACCCCTGACATCCTCTGTGGCCTTCGGCACTAATTTAGTCTGTATGCAGCCCCCTT	480
Db	141	SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu	160
Qy	481	AATCCGCTCTGCGCCTGCAGCAGCGGTACTAATCTCACATTATGGCCACAGAGCCTCC	540
Db	161	AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer	180
Qy	541	AATTATGCACAGTACCGGGTTGCCGCGCTACTATCCGTTACCGCCCTAGTGCTAAT	600
Db	181	AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn	200
Qy	601	GCAGTTCGAGGCTATGCTATATCCATTCTTCTGGCCTCAACACACACACCCCTACA	660
Db	201	AlaValGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr	220
Qy	661	TCTGTTGACATGAATCCATTACTTCCACTGATGTCAGGATTCTTGTTCAACCTGGCATA	720
Db	221	SerValAspMetAsnSerIleThrSerThrAspValAlaGlyLeuValGlnProGlyIle	240
Qy	721	GCATCTGAATTGTCATCCCAAGCGAGCGCCTTCACTACCGCAATCAAGTTGGCGCTCG	780
Db	241	AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer	260
Qy	781	GTTGAGACATCTGTGTGCTGAGAGGAAGCACTCCGCTCTGTCTATGCTATTATGCATA	840
Db	261	ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle	280
Qy	841	CATGGCTCTCCAGTTAACTCCTATACCAATACCCCTATACCGGTGCCCTTGCTTACTG	900
Db	281	HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu	300
Qy	901	GACTTTCCTTAGAGCTTGAGTTTCGCAATCTCACCACTGTAAACCAATACAGTGTG	960
Db	301	AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal	320

```
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 2 precursor (ORF2).
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=33774;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.J., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID. CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M80581; AAA45727.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW Signal.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70980 MW; 8085BC53CFB46FD3 CRC64;

Alignment Scores:
Pred. No.: 2.93e-139 Length: 660
Score: 149.00 Matches: 149
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.61% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST2_HEVPA (1-660)
QY 490 CTGCGCTGACGAGCGTACTACTACATTAATGCGCCAGAGGCTCCAATTATGCA 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisileMetAlaThrGluAlaSerAsnThrAla 183
QY 550 CAGTACCGGTTGCGCGGCTACTATCGTTACCGGCCCTAGTGCCTAATGCGATTGGA 609
Db 184 GlnTyrArgValAlaAlaGalaThrileArgTyrArgProLeuValProAsnAlaValGly 203
QY 610 GGCTATGCTATATCAATTCCTTCTGGGCTCAAAACCAACCAACCCCTACATCTGTGAC 669
Db 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCATTAATTCATGATGTCAGGATTCCTGTTCACCTGGCATGACATCTGAA 729
Db 224 MetAsnSerIleThrSerThrAspValArgileLeuValGlnProGlyIleAlaSerGlu 243
QY 730 TTGTCATCCCAAGGAGCGCTTCACTACCGCATCAAGTTGGCGCTCGGTGAGACA 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 263
QY 790 TCTGTTGTCTGAGGAGGAGCACCTCCGCTTGTGTCATGTTATGCATACATGGCTCT 849
Db 264 SerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283
QY 850 CCAGTTAACTCCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGCACTTTGCC 909
Db 284 ProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAla 303
```

```
QY 910 TTAGAGCTTCAGTTTCGCAATCTCACC 936
Db 304 LeuGluLeuGluPheArgAsnLeuThr 312
RESULT 3
VST2_HEVRH
ID VST2_HEVRH STANDARD; PRT; 485 AA.
AC Q00270;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Structural protein 2 (Fragment).
OS Hepatitis E virus (isolate Rhesus) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31766;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K.,
RA Min K.M.;
RT "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90274; BAA20910.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
FT NON TER 485 485
FT NON TER 1 1
SQ SEQUENCE 485 AA; 52317 MW; 5A4FD9273AC74F94 CRC64;

Alignment Scores:
Pred. No.: 2.14e-114 Length: 485
Score: 124.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.82% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST2_HEVRH (1-485)
QY 565 CGCGCTACTATCCGTTACCGGCCCTAGTGCCTAATGCGTTGAGGCTATGCTATATCC 624
Db 57 ArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSer 76
QY 625 ATTTCTTCTTGGCTCAACACCAACCCCTACATCTGTGACATGAATTCCTACTACT 684
Db 77 IleSerPheTrpProGlnThrThrThrProThrProThrSerValAspMetAsnSerIleThr 96
QY 685 TCCACTCATGTGAGGATTCCTTGTCAACCTGGCATAGCATCTGAATTTGGTCCATCCCAAGC 744
Db 97 SerThrAspValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSer 116
QY 745 GAGCGCTTCACTACCGCAATCAAGTTGGCGCTCGGTGAGACATCTGGTGTGCTGAG 804
Db 117 GluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGlu 136
QY 805 GAGGAAGCACCTCCGCTCTTGTTCATGTTATGCATACATGGCTCTCCAGTTAACTCTAT 864
Db 137 GluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyr 156
QY 865 ACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGGACTTTTGCCTTAGAGCTTGAGTTT 924
```

```
Db 157 ThrAsnThrProTyThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPhe 176
QY 925 CGCAATCTCACC 936
Db 177 ArgAsnLeuThr 180
RESULT 4
VST2_HEVMY STANDARD; PRT; 660 AA.
AC Q04611;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE Structural protein 2 precursor (ORF2)
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31769;
RN [1]
RP MEDLINE=93227573; PubMed=8470371;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
RL Virus Genes 7:95-109(1993).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10330; BAA01174.1; -.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW SIGNAL.
FT CHAIN 1 22 BY SIMILARITY.
FT CHAIN 23 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70998 MW; 3A82A4EA25C6253 CRC64;
Alignment Scores:
Pred. No.: 2,05e-114 Length: 660
Score: 124.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.82% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST2_HEVMY (1-660)
QY 565 CGCGTACTATCCGTACCGGCCCTAGTCCCTAATGAGTGGAGGCTATCTATATCC 624
Db 189 ArgAlaThrIleArgTyArgProLeuValProAsnAlaValGlyGlyTyrAlaIleSer 208
QY 625 ATTTCTTCTGCGCTCAACACACACACCCCTAGATCTGTGTGACATGAATCCATTACT 684
Db 209 IleSerPheTyrProGlnThrThrThrThrProThrSerValAspMetAsnSerIleThr 228
QY 685 TCCACTGATGTCAGATCTTCTTCAACTGGCATAGCATCTGAATGTGTCATCCCAAGC 744
Db 229 SerThrAspValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSer 248
QY 745 GAGCGCTTCACTACCGCAATCAAGTTGGCGCTCGGTGAGACATCTGGTCTGCTGAG 804
Db 249 GluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGlnThrSerGlyValAlaGlu 268
```

```
QY 805 GAGGAGCCACCTCCGGTCTCTGTCATGTTATGCATACATGGCTCTCCAGTTAACTCCTAT 864
Db 269 GluGluAlaThrSerGlyLeuValMetLeuCyHisGlySerProValAsnSerTyr 288
QY 865 ACCAATACCCCTTATATACCGGTGCCCTTGGCTTACTGGACTTTGCCCTTAGAGCTTGAAGTTT 924
Db 289 ThrAsnThrProTyThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPhe 308
QY 925 CGCAATCTCACC 936
Db 309 ArgAsnLeuThr 312
RESULT 5
VST2_HEVBV STANDARD; PRT; 660 AA.
AC F29326;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE Structural protein 2 precursor (ORF2)
OS Hepatitis E virus (strain Burma) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31767;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
CC -!- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA
CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M73218; AAA45736.1; -.
DR FIR; C40778; VHWNH2.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
KW SIGNAL.
FT CHAIN 1 19 POTENTIAL.
FT CHAIN 20 660 STRUCTURAL PROTEIN 2.
SQ SEQUENCE 660 AA; 70978 MW; 5832A013CCC4A61C CRC64;
Alignment Scores:
Pred. No.: 1.94e-110 Length: 660
Score: 120.00 Matches: 120
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.21% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST2_HEVBV (1-660)
```

```
QY 490 CTGCGCTGCGAGGACGGTACTAATACTACATTATGGCCACAGAGCGCTCCCAATTATGCA 549
Db 164 LeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSerAsnTyrAla 183
QY 550 CAGTACCGGTTGCCCGCGCTACTATCCGTTACCGGCCCTAGTGCCTAATCAGTTGGA 609
Db 184 GlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsnAlaValGly 203
```



```
QY 610 GGCTATGCTATCATCTTCTTCTGCGCTCAACCAACCAACCCCTACATCTGTTGAC 669
Db 204 GlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrProThrSerValAsp 223
QY 670 ATGAATTCCTACTTCCACTGATGTGAGGATCTTGTTCACCTGGCATGACATCTGAA 729
Db 224 MetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGlu 243
QY 730 TTGTCTATCCCAAGGAGCGCTTCACTACCGCAATCAAGTTGGCGCTCGGTTGAGACA 789
Db 244 LeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThr 263
QY 790 TCTCGTGTGCTGAGGAGGAGGACCTCCGCTCTTGTGTCATGTTATGATACATGGCTCT 849
Db 264 SerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySer 283

RESULT 6
VST1_HEVME
ID VST1_HEVME STANDARD; PRT; 123 AA.
AC Q03499, 1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Structural protein 1.
OS Hepatitis E virus (strain Mexico) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31768;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93079857; PubMed=1448913;
RA Huang C.C., Nguyen D., Fernandez J., Yun K.Y., Fry K.E.,
RA Bradley D.W., Tam A.W., Reyes G.R.;
RT "Molecular cloning and sequencing of the Mexico isolate of hepatitis
RT E virus (HEV).";
RL Virology 191:550-558(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74506; AAA45731.1; -.
DR PIR; C44212; C44212.
DR InterPro; IPR003384; HEV ORF2.
DR Pfam; PF02444; HEV ORF2; 1.
SQ SEQUENCE 123 AA; 12714 MW; C888F5D638852A68 CRC64;

Alignment Scores:
Pred. No.: 2,1e-99 Length: 123
Score: 109.00 Matches: 109
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.54% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST1_HEVME (1-123)

QY 2 TGCGCCCTAGGCTCTTTTGTGTTGTTTCTGCTTGTTCCTGCTATGTTGCCGGCCAC 61
Db 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGACGGGTGACGGCTCTGGCGCGCTGCTGGCGCGCGAGCGGCGGTACCGCGGTGTT 121
Db 35 ArgProValSerArgLeuAlaAlaValValGlyGlyAlaAlaAlaValProAlaValVal 54
QY 122 TCTGGGGTGACGGGTGATCTTCAGCCCTTGCCCAATCCCTATATTCATCCCAACCAACC 181
Db 55 SerGlyValThrGlyLeuIleLeuSerProSerGlnSerProIlePheIleGlnProThr 74
```

```
QY 182 CTTTGGCCAGAGCTTGGCGCTGCGTTCGACCTCGGCTTCGACCTCGGCTTCGCCAACAGGCC 241
Db 75 ProLeuProGlnThrLeuProLeuArgProGlyLeuAspLeuAlaPheAlaAsnGlnPro 94
QY 242 GGCCACTTGGCTCCACTTGGCGAGATCAGGCCAGGCCCTCCGCTGCTCCGCTGCC 301
Db 95 GlyHisLeuAlaProLeuGlyGluIleArgProSerAlaProLeuProValAla 114
QY 302 GACCTGCACAGCCGGCGCTGCGGCGC 328
Db 115 AspLeuProGlnProGlyLeuArgArg 123

RESULT 7
VST1_HEVBU
ID VST1_HEVBU STANDARD; PRT; 123 AA.
AC P29325;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Structural protein 1 (ORF3).
OS Hepatitis E virus (strain Burma) (HEV), and
OS Hepatitis E virus (strain Pakistan) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=31767, 33774;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Burma;
RA Tam A.W., Smith M.M., Guerra M.E., Huang M.E., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Pakistan;
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M73218; AAA45735.1; -.
DR PIR; M80581; AAA45726.1; -.
DR PIR; B40778; VHWHE.
DR InterPro; IPR003384; HEV ORF2.
DR Pfam; PF02444; HEV ORF2; 1.
SQ SEQUENCE 123 AA; 12676 MW; 8A5A798B1B74EDES CRC64;

Alignment Scores:
Pred. No.: 1,06e-51 Length: 123
Score: 61.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.26% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST1_HEVBU (1-123)

QY 2 TGCGCCCTAGGCTCTTTTGTGTTGTTTCTGCTTGTTCCTGCTATGTTGCCGGCCAC 61
Db 15 CysAlaLeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHis 34
QY 62 CGACGGGTGACGGCTCTGGCGCGCTGCTGGCGCGCGAGCGGCGGTACCGCGGTGTT 121
```

```
Db 35 ArgProValSerArgLeuAlaAlaValValGlyGlyAlaAlaValProAlaValVal 54
QY 122 TCTGGGTGACCGGTTGATTCCTCAGCCCTTCGCAATCCCTATATTATTCACCAACC 181
Db 55 SerGlyValThrGlyLeuLeuSerProSerGlnSerProIlePheIleGlnProThr 74
QY 182 CCT 184
Db 75 Pro 75
RESULT 8
VST1_HEVMY
ID VST1_HEVMY STANDARD; PRT; 123 AA.
AC Q04612;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Structural protein 1 (ORF3).
OS Hepatitis E virus (strain Myanmar) (HEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis B-like viruses.
OX NCBI_TaxID=31769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227573; PubMed=8470371;
RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,
RA Rikihisa T., Winn K.;
RT "Sequence and gene structure of the hepatitis E virus isolated from
RT Myanmar.";
RL Virus Genes 7:95-109 (1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10330; BAA01173.1; -
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
SQ SEQUENCE 123 AA; 12751 MW; FB81143F0B31F8A4 CRC64;
Alignment Scores:
Pred. No.: Length: 123
Score: 59.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.95% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x VST1_HEVMY (1-123)
QY 8 CTAGCCCTCTTTGCTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 67
Db 17 LeuGlyLeuPheCysCysSerSerCysPheCysLeuCysCysProArgHisArgPro 36
QY 68 GTACGCCCTCTGGCGCGCTGTCGGCGCGCGAGCGCGGTACCGCGGTGTTCTGGG 127
Db 37 ValSerArgLeuAlaAlaValValGlyGlyAlaAlaValProAlaValValSerGly 56
QY 128 GTACCGCGGTGATTCCTCAGCCCTTCGCAATCCCTATATTATTCACCAACCCT 184
Db 57 ValThrGlyLeuLeuSerProSerGlnSerProIlePheIleGlnProThrPro 75
RESULT 9
ACCA_BACSU
ID ACCA_BACSU STANDARD; PRT; 325 AA.
AC O34847;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
```

```
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha
DE (EC 6.4.1.2)
GN ACCA OR BS029200.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RT in the 200 kb rrmB-dnaB region.";
RL Microbiology 143:3431-3441 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Fezzari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerz-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinola S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nockack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256 (1997).
CC -!- FUNCTION: This protein is a component of the acetyl coenzyme A
CC carboxylase complex; first, biotin carboxylase catalyzes the
CC carboxylation of the carrier protein and then the transcarboxylase
CC transfers the carboxyl group to form malonyl-CoA (by similarity).
CC -!- CATALYTIC ACTIVITY: Carboxybiotin carboxyl carrier protein +
CC acetyl-CoA = biotin carboxyl carrier protein + malonyl-CoA.
CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX (BY SIMILARITY).
CC -!- SIMILARITY: TO THE C-TERMINUS OF MAMMALIAN PROPIONYL-COA
CC CARBOXYLASE BETA CHAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF008220; AAC00341.1; -
DR EMBL; Z99118; CAB14880.1; -
```

```
DR PIR; G69580; G69580.
DR Subtilisin; BGI2557; accA.
DR InterPro; IPR001095; Ac-CoA_carboxylA.
DR Pfam; PF03255; ACCA; 1.
DR PRINTS; PR01069; ACCTRFRASEA.
DR TIGRFAMS; TIGR00513; accA; 1.
KW Fatty acid biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 325 AA; 36333 MW; 9B177DEB4A5B5864 CRC64;

Alignment Scores:
Pred. No.: 4.42 Length: 325
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.40% Indels: 0
DB: Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x ACCA_BACSU (1-325)
QY 294 GAGGACGAGCGGCGGCTGGGCTG 268
DB 199 GlyGlySerGlyGlyAlaLeuGlyLeu 207

RESULT 10
RL7_CLOPE STANDARD; PRT; 121 AA.
AC Q8XHR7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12.
GN RPLL OR CPE2414.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -!- FUNCTION: Seems to be the binding site for several of the factors
CC involved in protein synthesis and appears to be essential for
CC accurate translation (by similarity).
CC -!- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP003194; BAB82120.1; -.
CC HAMAP; MF_00368; -; 1.
DR InterPro; IPR008932; Ribos L12/7 olig.
DR InterPro; IPR002026; Ribosomal L12.
DR Pfam; PF00542; Ribosomal L12; 1.
DR ProDom; PD001326; Ribosomal_L12; 1.
DR TIGRFAMS; TIGR00855; L12; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 121 AA; 12535 MW; 5AF9F83D2C94AC3E CRC64;

Alignment Scores:
Pred. No.: 50.1 Length: 121
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.21% Indels: 0
DB: Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x RL7_CLOPE (1-121)
QY 83 GCCGTCGTCGGCGCGCGAGCGGCG 106
DB 38 AlaValValGlyGlyAlaAlaAla 45

RESULT 11
E315_ADE05 STANDARD; PRT; 132 AA.
ID E315_ADE05
AC P06498;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Early E3B 14.6 kDa protein precursor.
OS Human adenovirus type 5.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28285;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85092388; PubMed=2981456;
RA Cladaras C., Wold W.S.M.;
RT "DNA sequence of the early E3 transcription unit of adenovirus 5.";
RL Virology 140:28-43(1985).
RN [2]
RP IDENTIFICATION OF PROTEIN
RX MEDLINE=90177214; PubMed=2309441;
RA Tollefson A.E., Krajcsi P., Pursley M.H., Gooding L.R., Wold W.S.M.;
RT "A 14,500 MW protein is coded by region E3 of group C human
RT adenoviruses.";
RL Virology 175:19-29(1990).
RN [3]
RP SIGNAL SEQUENCE CLEAVAGE SITE.
RX MEDLINE=92148953; PubMed=1531370;
RA Krajcsi P., Tollefson A.E., Anderson C.W., Wold W.S.M.;
RT "The adenovirus E3 14.5-kilodalton protein, which is required for
RT down-regulation of the epidermal growth factor receptor and
RT prevention of tumor necrosis factor cytotoxicity, is an integral
RT membrane protein oriented with its C terminus in the cytoplasm.";
RL J. Virol. 66:1665-1673(1992).
RN [4]
RP PHOSPHORYLATION.
RX MEDLINE=92189523; PubMed=1532104;
RA Krajcsi P., Wold W.S.M.;
RT "The adenovirus E3-14.5K protein which is required for prevention of
RT TNF cytotoxicity and for down-regulation of the EGF receptor contains
RT phosphoserine.";
RL Virology 187:492-498(1992).
RN [5]
RP O-GLYCOSYLATION.
RX MEDLINE=92263760; PubMed=1533979;
RA Krajcsi P., Tollefson A.E., Wold W.S.M.;
RT "The E3-14.5K integral membrane protein of adenovirus that is
RT required for down-regulation of the EGF receptor and for prevention
RT of TNF cytotoxicity is O-glycosylated but not N-glycosylated.";
RL Virology 188:570-579(1992).
RN [6]
RP COMPLETE GENOME.
RX MEDLINE=92087470; PubMed=1727603;
RA Chroboczek J., Bieber F., Jacrot B.;
RT "The sequence of the genome of adenovirus type 5 and its comparison
RT with the genome of adenovirus type 2.";
RL Virology 186:280-285(1992).
CC -!- FUNCTION: Down-regulates the EGF receptor and prevents cytotoxicity
CC by TNF.
CC -!- PTM: Phosphorylated on serine; O-glycosylated, but not N-
CC glycosylated.
CC -!- SIMILARITY: TO EQUIVALENT PROTEIN IN OTHER ADENOVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; M73260; ; NOT ANNOTATED_CDS.
DR EMBL; X03002; CAA26786.1; -
DR InterPro: IPR008131; Adeno_E3_14_5.
DR Pfam; PF04834; Adeno_E3_14_5; 1.
KW Early protein; Transmembrane; Glycoprotein; Phosphorylation; Signal.
FT SIGNAL 1 18
FT CHAIN 19 132 EARLY E3B 14.6 kDa PROTEIN.
FT TRANSMEM 51 75 POTENTIAL.
SQ SEQUENCE 132 AA; 14750 MW; 524690C4AD9B9A74 CRC64;

Alignment Scores:
Pred. No.: 49.5 Length: 132
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.24% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x E315_ADE05 (1-132)

QY 1097 CCGCGCGGACCTTCACCTACCCCA 1074

Db 112 ProArgProThrSerProthrPro 119

RESULT 12

SORBP

ID SORBP PIG STANDARD; PRT; 153 AA.

AC P28220;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Sorbin.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE.

RC TISSUE=Intestine;

RX MEDLINE=92007891; PubMed=1915377;

RA Vagne-Descroix M., Pansu D., Joernvall H., Carlquist M., Guignard H.,

RA Jourdan G., Desvigne A., Collinet M., Callet C., Mutt V.,

RT "Isolation and characterisation of porcine sorbin";

RL Eur. J. Biochem. 201:53-59(1991).

CC -!- FUNCTION: Increases water and sodium absorption in the intestine

CC and gall-bladder.

CC -!- SIMILARITY: Contains 1 SoHo domain.

DR PIR; S17837; S17837.

DR InterPro; IPR003127; Sorb.

DR Pfam; PF02208; Sorb; 1.

DR ProDom; PD016158; Sorb; 1.

DR SMART; SM00459; Sorb; 1.

DR PROSITE; PS50831; SOHO; 1.

FT DOMAIN 1 46 SOHO.

SQ SEQUENCE 153 AA; 17481 MW; A49C219455BB1417 CRC64;

Alignment Scores:

Pred. No.: 48.5 Length: 153

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.24% Indels: 0

DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x SORBP_PIG (1-153)

QY 119 CCACCGCGGTACCGCGTGCGC 96

Db 93 ProProProValProProLeuArg 100

RESULT 13

UL03_HSV2H

ID UL03_HSV2H STANDARD; PRT; 233 AA.

AC P28279;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 01-DEC-1992 (Rel. 24, Last annotation update)

DE Protein UL3.

GN UL3.

OS Herpes simplex virus (type 2 / strain HG52).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Simplexvirus.

OX NCBI_TaxID=10315;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92113549; PubMed=1662697;

RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;

RT "Comparative sequence analysis of the long repeat regions and

RT adjoining parts of the long unique regions in the genomes of herpes

RT simplex viruses types 1 and 2";

RL J. Gen. Virol. 72:3057-3075(1991).

CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL3,

CC HSV-2 UL3, EBV-1 60, AND VZV 58.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D10470; BAA01266.1; -.

DR PIR; JQ1496; WMBEHL.

DR InterPro; IPR005035; Herpes UL3.

DR Pfam; PF03369; Herpes UL3; 1.

SQ SEQUENCE 233 AA; 25649 MW; B38F613BD839AA24 CRC64;

Alignment Scores:

Pred. No.: 45.7 Length: 233

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.21% Indels: 0

DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x UL03_HSV2H (1-233)

QY 421 TCACCCCTGACATCCTCTGTGCC 444

Db 75 SerProLeuThrSerSerValala 82

RESULT 14

R22_PORPU

ID R22_PORPU STANDARD; PRT; 241 AA.

AC P51249;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Chloroplast 30S ribosomal protein S2.

GN RPS2.

OS Porphyra purpurea.

OG Chloroplast.

OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Porphyra.

OX NCBI_TaxID=2787;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Avonport;

RA Reith M.E., Munnholland J.;

RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 16, 2004, 13:41:04 ; Search time 19.9849 Seconds
(without alignments)
10229.686 Million cell updates/sec

Title: US-09-851-410A-10_COPY_5117_7096

Perfect score: 659

Sequence: 1 ATGGCGCCTAGGCTCTTTT.....GTAAACTCGGAGTTCGTAG 1980

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Word size: 1

Total number of hits satisfying chosen parameters: 663654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p_model -DEV=xlp
-Q=/cgn2_1/USPRO_spool_p/US09851410/runat_16082004_125718_9682/app_query.fasta_1.4238
-DB=Issued Patents AA -QMT=fastan -SUFFIX=oli.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09851410@cgn1_1_45@runat_16082004_125718_9682 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pap:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pap:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pap:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pap:*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pap:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	659	100.0	659	1	US-08-240-049B-16
2	659	100.0	659	1	US-08-259-148A-20
3	659	100.0	659	1	US-08-484-054-20
4	659	100.0	659	2	US-07-876-941A-20
5	659	100.0	659	3	US-08-477-292-14
6	659	100.0	659	4	US-07-870-985A-20
7	548	83.2	548	3	US-08-477-292-14
8	512	77.7	660	3	US-08-542-634-14
9	512	77.7	660	5	PCT-US95-13703-14
10	438	66.5	660	4	US-09-462-606-49
11	435	66.0	435	1	US-08-259-148A-18
12	435	66.0	435	1	US-08-484-054-18

13	435	66.0	435	2	US-07-876-941A-18	Sequence 18, Appl
14	435	66.0	435	4	US-07-870-985A-18	Sequence 18, Appl
15	401	60.8	549	3	US-08-542-634-16	Sequence 16, Appl
16	401	60.8	549	5	PCT-US95-13703-16	Sequence 16, Appl
17	392	59.5	540	3	US-08-542-634-26	Sequence 26, Appl
18	392	59.5	540	5	PCT-US95-13703-26	Sequence 26, Appl
19	377	57.2	525	3	US-08-542-634-28	Sequence 28, Appl
20	377	57.2	525	5	PCT-US95-13703-28	Sequence 28, Appl
21	327	49.6	327	1	US-08-240-049B-14	Sequence 14, Appl
22	327	49.6	327	1	US-08-259-148A-16	Sequence 16, Appl
23	327	49.6	327	1	US-08-484-054-16	Sequence 16, Appl
24	327	49.6	327	2	US-07-876-941A-16	Sequence 16, Appl
25	327	49.6	327	2	US-08-542-634-18	Sequence 18, Appl
26	327	49.6	327	3	US-08-477-292-18	Sequence 18, Appl
27	327	49.6	327	4	US-07-870-985A-16	Sequence 16, Appl
28	327	49.6	327	5	PCT-US95-13703-18	Sequence 18, Appl
29	149	22.6	552	4	US-03-172-699-16	Sequence 16, Appl
30	149	22.6	561	4	US-09-172-699-20	Sequence 20, Appl
31	149	22.6	660	4	US-09-462-606-12	Sequence 12, Appl
32	149	22.6	660	4	US-09-462-606-49	Sequence 49, Appl
33	149	22.6	660	4	US-09-462-606-51	Sequence 51, Appl
34	149	22.6	660	4	US-09-172-699-2	Sequence 2, Appl
35	124	18.8	660	4	US-09-462-606-55	Sequence 55, Appl
36	120	18.2	525	3	US-08-542-634-27	Sequence 27, Appl
37	120	18.2	525	5	PCT-US95-13703-27	Sequence 27, Appl
38	120	18.2	540	3	US-08-542-634-25	Sequence 25, Appl
39	120	18.2	540	5	PCT-US95-13703-25	Sequence 25, Appl
40	120	18.2	549	3	US-08-542-634-15	Sequence 15, Appl
41	120	18.2	549	3	US-08-477-292-15	Sequence 15, Appl
42	120	18.2	549	5	PCT-US95-13703-15	Sequence 15, Appl
43	120	18.2	660	1	US-08-240-049B-15	Sequence 15, Appl
44	120	18.2	660	1	US-08-259-148A-19	Sequence 19, Appl
45	120	18.2	660	1	US-08-484-054-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1

US-08-240-049B-16
; Sequence 16, Application US/08240049B
; Patent No. 5686239
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Tam, Albert W.
; APPLICANT: Yarbough, Patrice O.
; TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,049B
; FILING DATE: 09-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles K. Sholtz
; REGISTRATION NUMBER: 38, 615
; REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)
; INDIVIDUAL ISOLATE: ORF-2
US-08-240-049B-16

Alignment Scores:
Pred. No.:      0      Length:      659
Score:          659.00    Matches:      659
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match:      100.00%    Indels:      0
DB:                1      Gaps:      0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-08-240-049B-16 (1-659)
QY 1 ATGCGCCCTAGGCTCTTTTGTGTTGTTCTCTTGTGTTCTGCTATGTCGCGGCCA 60
DB 1 MetArgProArgProLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CCGACCGGTACGCGTCTCGCGCCCGTCTGCGCGCGCGGCGGTACCGCGGTGT 120
DB 21 ProThrGlyGlnProSerGlyArgArgGlyArgSerGlyGlyThrGlyGly 40
QY 121 TTCTGGGTGACGGTGTGATCTCAGCCCTTCGCATCCCTATATTCATCAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTrpTrpHisProThrAsn 60
QY 181 CCCTTTGCCCCAGAGCTTCGGCTGCTCGGCTCTGGACCTCGCTCGCCCAACAGCC 240
DB 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGGCCACTTGCTCCACTTGGCGAGATCAGGCCCGCCCTCGCTCCGCTCCGCTCGC 300
DB 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArg 100
QY 301 CGACCTGCCACAGCGGGCTGGCGCTGAGCGTGTGGCGCTGCGCCATCAGACCTCA 360
DB 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCGCTCCCGGACGTTCATCTCGCGTGCAATTTCTACGCGCGCAGTATAATTGTCTACT 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTrpAsnLeuSerThr 140
QY 421 TCACCCCTGACATCCTCTGTGGCCTCTGGCCTAAATTTAGTCTCTGTATGACGCCCTT 480
DB 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AATCCGCTCTCGCGTGCAGACGGTACTAATCTACATTCATGCGCCACAGAGGCTCC 540
DB 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGACAGTACCGGGTTCGCGCTACTATCTCCGTTACCGGCCCTAGTGCCTAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCAGTTGGAGGTATGCTATATCCATTTCTTTCTGGCCCTCAACCAACCAACCCCTACA 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
QY 661 TCTGTTGATGATGAATCCATTAATCTCCACTGATGTCAGGATCTTGTTCACCTGGCATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATTGGTTCATCCACGAGCGCCTTCTACTACCGCAATCAAGGTTGGGCTCG 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTAGACATCTGGTGTCTGAGGAGGAGCCACTCCGGTCTTGTGTCATGTTATGATGATA 840
DB
```

```
DB 261 ValGluThrSerGlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCTCCAGTTAACTCTATACCAATAACCCCTTATACCGGTGCCCTTGGCTTACTG 900
DB 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTTGCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAAACCAACCAATACACGTG 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTACTCCAGCACTCTCTGTCACCTCCGCCCGAGGGCGGACGGGACTGCGGAGCTG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCAACAAGTGCAGCCACAGGTTTCATGAAGATCTCCACTTTACCGGCTTTAATGGGTA 1080
DB 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAAGTCGCGCGGGATAGCTCTAACATTACTTAACCTTGTCTGACACGCTCCTCGGC 1140
DB 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGCTCCCGACAGAATAATTTCTGCTGCTGCGCGGCACTGTTTTATTCCCGCCCGGT 1200
DB 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GTCTCAGCCAATGGCGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAACTGCTCAGCAG 1260
DB 401 ValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAAGGTGTGCTATCCCCACCATATCGATCTTGGTGATTGGCTGTGGTCAATTCAG 1320
DB 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGln 440
QY 1321 GATTATGACACACGATGAGCAGGATCGGCCACCCCGTCCGCTCGCCATCTCGGCT 1380
DB 441 AspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTCTCCGAGCAATATGATCTTTGGCTGCTCCCTCCTCCTCAGCCGCGGATGAC 1440
DB 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAsp 480
QY 1441 CAGTCCACTTACGGTCTGCTCAACTGGCCCGGTTTATATCTCGACAGCGCTGCTTTGGTG 1500
DB 481 GlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTGCACTGGCGCGCAGCCCTAGCCGATCGCTTGACTGTGCCAAGTCAACCTC 1560
DB 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
QY 1561 GACGGCGGCCCTCCGACTGTTGAGCAATATTCAGACATTTCTTGTGCTCCGCTT 1620
DB 521 AspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGCAAGCTCTCCTTTTGGAGCGCCGACAAAGAGGTTATCTCTATAATAT 1680
DB 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyr 560
QY 1681 AATACTACTGTAGTACAGATTCTGATTGAAATGTGCGGCCATCGGTCGCCAAT 1740
DB 561 AsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580
QY 1741 TCAACCTATACCACAGGCTGGGCGCGTCCGCTCGCATTTCTGCGCGCGGCTTTG 1800
DB 581 SerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeu 600
QY 1801 GCTCCACGCTCCGCGCTCTGCTGTGAGGATATCTTTGATTATCCGGGGCGGCGCAC 1860
DB 601 AlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHis 620
QY 1861 ACATTTGATGATCTTCCCTCGCGCTGAATGCGCGCTTTAGGCTCCAGGTTGTGCTTCCAG 1920
DB 621 ThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGln 640
```


Db 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCCAACTCGACGACACAGGTTCTAGAAAGATCTCCACTTTACCGCCCTTAATGGGTA 1080
Db 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAAGTCGGCGGGGATAGCTCTAACATTACTTTAACTTGTGACACGCTCCCTCGGC 1140
Db 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGCTCCGACAGAATTAAATTTGCTGGCTGGCGGCAACTGTTTATTCGCCCGCGTT 1200
Db 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GTCCTCAGCAATGGCGACCAACCTGAGCTCTATACATGAGTGGAGAACTCTCAGCAG 1260
Db 401 ValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAAGGGTGTGCTATCCCCACGATATCGATCTTGGTGATTGCGGTGTGTCATTTCAG 1320
Db 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValIleGln 440
QY 1321 GATTATGACAACAGCATGAGCAGGATCGGCCACCCCTCGCTCGCTGCGCATCTCGGCT 1380
Db 441 AspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTTCTCCGAGCAAAATGATCTACTTTGGCTGCTCCCTCACTCAGCCGAGTATGAC 1440
Db 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAsp 480
QY 1441 CAGTCCACTTACGGTCTGCTCAACTGGCGCGTTTATATCTCGGACAGCGTACTTTGGTG 1500
Db 481 GlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTGCACTGGCGGAGCCGCTAGCCCGATCGTTGACTGGTCCAAAGTCACCCCTC 1560
Db 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
QY 1561 GACGGCGGCCCTCCCGACTGTTCCAGCAATATCCAGACATCTTTGTGCTCCCCCTT 1620
Db 521 AspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGCAAGCTCTCTCTTTGGAGCGCGGCACACAAAAGCAGGTTATCTCTTATAATTAT 1680
Db 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyr 560
QY 1681 AATACTACTGAGTGACAGATTCTGATTGAAATGCTGCGGCCATCGGCTCGCCATT 1740
Db 561 AsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580
QY 1741 TCAACTATACCAACAGGCTGGGCGCGTCCGGTCCGCTTCTGCGCGCGGTTTG 1800
Db 581 SerThrTyrThrThraArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeu 600
QY 1801 GCTCCAGCTCGGCTCGCTCTGCTGGAGGATCTTTTGATTATCCGGGCGGCGCAC 1860
Db 601 AlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHis 620
QY 1861 ACATTGTGACTTCTGCGCTGAATGCGCGCTTTAGGCTCCAGGTTGCTTTCAG 1920
Db 621 ThrPheAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGln 640
QY 1921 TCAACTGTCGCTAGCTCAGCGCTTAAAGTTAAGTGGGTAAACCTCGGGAGTTG 1977
Db 641 SerThrValAlaGluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 659

RESULT 3

US-08-484-054-20
; Sequence 20, Application US/08484054
; Patent No. 5770689
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.

; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,054
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 870,985
; FILING DATE: 20-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 9
US-08-484-054-20

Alignment Scores:
Pred. No.: 0 Length: 659
Score: 559.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-08-484-054-20 (1-659)

QY 1 ATGCGCCCTAGGCTCTTTTGTGCTGTTCCTCTGTTGTTGCTGCTATGTTGCCGGGCCA 60

[illegible]

361	GlyC	Val	Gly	Arg	Gly	Ile	Ala	Leu	Thr	Leu	Leu	Asn	Leu	Ala	Asp	Thr	Leu	Leu	Gly	380		
1141	GGG	CTC	CCC	GAC	AGA	TTA	ATTT	TCG	TGG	CGG	CGG	CAAC	TGTTTT	ATTT	CCCG	CCCG	GGT	1200				
381	Gly	Leu	Pro	Thr	Glu	Leu	Ile	Ser	Ser	Ala	Gly	Gly	Ile	Leu	Phe	Tyr	Ser	Arg	Pro	Val	400	
1201	GTCT	CAG	CCA	TAT	CGC	GAC	CAA	CCG	TGA	AGCT	CTA	TAC	TAC	TAC	TAC	TAC	TAC	TAC	TAC	TAC	1260	
401	Val	Ser	Ala	Asn	Gly	Glu	Pro	Thr	Val	Leu	Tyr	Thr	Ser	Val	Glu	Asn	Ala	Gln	Ile	420		
1261	GATA	AA	GGG	TGT	TGCT	ATC	CCCC	CCAC	GAT	ATCC	ATCT	TGGT	GAT	TTC	GGCT	TGGT	TGGT	CA	TT	TCAG	1320	
421	Asp	Lys	Gly	Val	Ala	Ile	Pro	His	Asp	Ile	Asp	Lys	Gly	Asp	Ser	Arg	Val	Val	Ile	Gln	440	
1321	GATT	AT	GAC	AA	CAC	GAT	ATG	ACG	AGG	ATCG	CGC	CAC	CCCC	TGC	GCT	CGC	CA	CT	CTC	GGC	CT	1380
441	Asp	Tyr	Asp	Asn	Gln	His	Glu	Asp	Arg	Pro	Thr	Ser	Pro	Ala	Pro	Ser	Arg	Pro	460			
1381	TTTT	CTCT	CTT	CCG	AG	CAA	TGAT	GTACT	TTT	GGCT	GTCC	CTCAT	CTG	CAG	CGC	GAC	TAT	GCAC	1440			
461	Phe	Ser	Val	Leu	Arg	Ala	Asn	Asp	Val	Leu	Trp	Leu	Ser	Leu	Thr	Ala	Glu	Tyr	Asp	480		
1441	CAGT	CCA	CTT	TAC	GGG	TGCT	CAAC	TGG	CCCC	GGTTT	TAT	TCT	CGG	AC	ACG	CTG	CA	TTT	GGTG	1500		
481	Gln	Ser	Thr	Tyr	Gly	Ser	Ser	Thr	Gly	Pro	Val	Tyr	Ile	Ser	Asp	Ser	Val	Thr	Leu	Val	500	
1501	AATG	TT	TCG	CACT	GCG	CGC	GAG	CGG	TAC	CGG	ATCG	CTG	CTG	CACT	GGT	CCCA	AGT	CA	CCCTC	1560		
501	Asn	Val	Ala	Thr	Gly	Ala	Gln	Ala	Val	Ala	Arg	Ser	Leu	Asp	Trp	Ser	Lys	Val	Thr	Leu	520	
1561	GACG	GGG	GGC	CCCT	CCCC	AGT	GTG	TGAC	CAAT	ATAT	TCCA	AGAC	ATAT	TCT	TGT	TGCT	CC	CCGCTT	1620			
521	Asp	Gly	Arg	Pro	Leu	Pro	Thr	Val	Glu	Gln	Tyr	Ser	Lys	Thr	Phe	Phe	Val	Leu	Pro	Leu	540	
1621	CTGG	CAAG	CTCT	CTCT	TTT	CGG	AGC	CGG	CAC	AA	CAAA	AG	CAG	GT	TAT	CC	T	T	A	T	1680	
541	Arg	Gly	Lys	Leu	Ser	Phe	Trp	Glu	Ala	Gly	Thr	Thr	Lys	Ala	Gly	Tyr	Pro	Tyr	Asn	Tyr	560	
1681	AAT	TACT	ACT	CT	TAG	TGAC	CA	GAT	TC	GAT	TG	AAAT	CT	CGC	GGC	CA	T	CGG	T	CGC	CA	1740
561	Asn	Thr	Thr	Ala	Ser	Asp	Gln	Ile	Ile	u	leu	Ile	u	leu	Ala	Gly	His	Arg	Val	Ala	Ile	580
1741	TC	AACT	AT	TAC	CA	CC	AGG	CTT	G	GGG	CGC	GT	CGG	CT	CGC	CA	TTT	CT	CGG	CGC	GGT	1800
581	Ser	Thr	Tyr	Thr	Thr	Arg	Leu	Gly	Ala	Gly	Pro	Val	Ala	Ile	Ser	Ala	Ala	Val	Leu	600		
1801	GCT	CC	AGCT	CCG	CCCT	GGCT	CTCT	GCT	CG	AGG	AT	ACT	TTT	TG	AT	TAT	CCG	GGC	GGC	GGC	CA	1860
601	Ala	Pro	Arg	Ser	Ala	Leu	Ala	Leu	Leu	Glu	Asp	Thr	Phe	Asp	Tyr	Pro	Gly	Arg	Ala	His	620	
1861	ACA	TTT	CGAT	CAC	TCT	CT	CCCT	TGA	T	CGC	GGC	CTT	TAG	CGCT	TCC	ACG	GGT	TGT	GTCT	TC	CA	1920
621	Thr	Phe	Asp	Asp	Phe	Cys	Pro	Glu	Cys	Arg	Ala	Leu	Gly	Leu	Gln	Gly	Cys	Ala	Phe	Gln	640	
1921	TC	AACT	GT	CGCT	GAC	TGCT	CC	AGC	CGCT	TAA	AGT	TAA	GGT	TGG	TG	TAA	CT	CGG				

RESULT 4
US-07-876-941A-20
; Sequence 20, Application US/07876941A
; Patent No. 5885768
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Tam, Albert W.
; APPLICANT: Mitchell, Carl
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and
; ANTI-BODIES
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/876,941A
FILING DATE: 01-MAY-1992
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 822,335
FILING DATE: 17-JAN-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 505,888
FILING DATE: 05-APRIL-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 420,921
FILING DATE: 13-OCTOBER-1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 367,486
FILING DATE: 16-JUNE-1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 336,672
FILING DATE: 11-APRIL-1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 208,997
FILING DATE: 17-JUNE-1988
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0093.33
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 9
US-07-876-941A-20

Alignment Scores:
Pred. No.: 0 Length: 659
Score: 659.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-07-876-941A-20 (1-659)

QY 1 ATGGCCCTAGGCTCTTTGCTGTTGCTCTTCTGTTCTGCTATGTTGCCCGCGCA 60
DB 1 MetArgProArgProLeuLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CGACCGGTACGCGCTCTGCGCGCGCTCTGTTGGCGCGCGAGCGGCTACCGCGGTGT 120
DB 21 ProThrGlyGlnProSerGlyArgArgArgGlyArgArgSerGlyGlyThrGlyGly 40
QY 121 TTCGGGTGACCGGGTGAATTCCTCAGCCCTTCGAAATCCCTATATCATCAACCAAC 180
DB 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyriIleHisProThrAsn 60
QY 181 CCCTTGGCCACAGCGTTCGCGCTCGCGCTCGGACCTCGGCTTCGCCAACCGCC 240

61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGGCCACTTGGCTCCACTTGGCGAGATCAGGCCCGCCCGCTCCGCTCCGCTCCGCTCCG 300
DB 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArg 100
QY 301 CGACCTGCCACAGCGGGGCTCGCGCGCTGACGGCTGTGGCGCTCCGCCATCACACTCA 360
DB 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCCGTCGGGACGTTGATTCCTCGCGTGCAATTCACGCCGACGATTAATTTGTCTACT 420
DB 121 ProValProAspValAspSerArgGlyAlaIleLeuArgArgGlnTyrAsnLeuSerThr 140
QY 421 TCACCCCTGACATCCTCTGTGGCTCTGSCACTAATTTAGTCTCTGTATGCACGCCCTT 480
DB 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AATCCGCTCTCGCGCTCGCAGACGGTACTAATATCTACATTTATGGCCACAGAGGCTCC 540
DB 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGACAGTACCGGGTTCGCCGCTACTATCCGTTACCGGCGCTCTAGTGCCTAAT 600
DB 181 AsnTyrAlaGlnTyrArgValAlaAlaAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTCGCCCTCAACCAACCAACCCCTACA 660
DB 201 AlaValGlyGlyTyrAlaIleSerIleSerPheTyrProGlnThrThrThrProThr 220
QY 661 TCTGTTGACATGAATTCCTACTTCCACTGATGTGAGGATTTCTGTTCAACCTGGCATA 720
DB 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATTCGTCATCCAGAGCGGCTTACTACCGCAATCAAGTTGGGCTCG 780
DB 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGGAGGACCTCCGCTCTGTCATGTTATGATA 840
DB 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGCTCTCCAGTTAACTCTCTATACCAATACCCCTTATACCGTCCCTCTGGCTACTG 900
DB 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTGTGCTTAGACCTTGAGTTTCGCAATCTCACACCTGTAAACCAATACAGTGTG 960
DB 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTACTCCAGCACTGCTGTCACCTCCGCGCGGCGCGAGGACTCGGAGCTG 1020
DB 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCCAACTGCAGCCACCGAGGTTTCATGAAAGATCTCCACTTTACCGGCTTAAATGGGTA 1080
DB 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAGTCCGCGCGGATAGCTTAACATTAATTAACCTTGCTGACACGCTCCCTCGC 1140
DB 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGCTCCCGACAGAAATTAATTTGCTGCGCTGGCGGCAACTGTTTATTCGCGCGCTT 1200
DB 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GTCTCAGCCAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAGTCTCAGCAG 1260
DB 401 ValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAGGCTGTGCTATCCCGACGATATCCATCTTGTGATTCCGCTGTGCTCATTCAG 1320
DB 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGln 440


```
Db 201 AlaValGlyGlyTyraAlaIleSerIleSerPheTrpProGlnThrThrThrThrProThr 220
QY 661 TCTGTTGACATGAATTCCTACTTCCACTGATGTCAGGATTCCTGTTCAACCTGCAT 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleuValGlnProGlyIle 240
QY 721 GCATCTGAATTTGGTCATCCCAAGCAGCGCCCTTCACTACCGCAATCAAGGTTGGGCTCG 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTGAGACATCTGGTGTCTCTGAGGAGGAAGCACCTCCGGTCTGTGTCATGTATGCATA 840
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCTCCAGTTAACTCCATACCAATACCCCTTATACCGGTGCGCTTGGCTTACTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGGCTTAGAGCTTGAGTTTCGCAATCTCACACCTGTAACACCAATACAGTGTG 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTCTACTCCAGCAGCTCTCGTCACTCCGCCGAGGGCCGAGCGGACTCGGAGCTG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCCAACTGCAGCCACCAAGGTTTCATGAAGATCTCCACTTTACCGGCTTAATGGGTA 1080
Db 341 ThrThrThrAlaAlaThrArgPheMetIysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GTGAAGTCGGCGGGGATAGCTCTACATTACTTAACCTTGCTGACACGCTCTCTCGGC 1140
Db 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGCTCCGACAGAAATTAATTTCTGTCGGTGGGGGCACTGTTTATTCGGCGCGGTT 1200
Db 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GTCTCAGCCAAATGGCGAGCAACCGTGAAAGCTCTATACATCAGTGGAGAACTCTCAGAG 1260
Db 401 ValSerAlaAsnGlyGluProThrValIysLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAGAGGTGTTGCTATCCCCACAGATPCGATCTTGTGTGATTCGCGTGTGTCATTACAG 1320
Db 421 AspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValIleGln 440
QY 1321 GATTATGACACACGATGACGAGATCGGCCACCCGTCGCTCGCTCGCCATCTCGGCT 1380
Db 441 AspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTTCTCCGAGCAAAATGATGTACTTTGGCTGTCCCTCACCTGCAGCCGAGTATGAC 1440
Db 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAsp 480
QY 1441 CAGTCCACTTAGCGTGTCAACTGGCCGGTTHATATCTCGGACAGCGTGTACTTTGGTG 1500
Db 481 GlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTGCGACTGGCGCGCAGCCGTAGCCGATCGCTTGCTGCTGCTCAAGTCAACCTC 1560
Db 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
QY 1561 GACGGCGGCGCCCTCCCGACTGTTGAGCAATATTCAGCAATCTTTGTGCTCCCGCTT 1620
Db 521 AspGlyArgProLeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGCAAGCTCTCTTTTGGAGCGCGGCACAAACAAAGCAGGTTATCTTTATTAATAT 1680
Db 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyr 560
QY 1681 AATACTACTGCTAGTACACAGATTCCTGATTAAGAAATGCTGCGCGCATCGGGTCCCAT 1740
Db 561 AsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580
```

```
QY 1741 TCAACCTATACACAGGCTTCGGGCGCGTTCGGTCCGTCGCATTTCTCGGCGCGGCTTTG 1800
Db 581 SerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeu 600
QY 1801 GTCACACGCTCCGCCCTCGCTCTGCTGAGAGGATACCTTTTGATTTATCCGGGCGGCGCAC 1860
Db 601 AlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHis 620
QY 1861 ACATTTGATGACTTTCGCCCTGAAATCGCCGCTTTAGGCTCCAGGTTGTGCTTTCCAG 1920
Db 621 ThrPheAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGln 640
QY 1921 TCAACTGTCTGAGCTCCAGCGCTTAAAGTTAAAGTGGTAAACTCGGAGTTG 1977
Db 641 SerThrValAlaGluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 659
```

RESULT 6
US-07-870-985A-20
; Sequence 20, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/870,985A
; FILING DATE: 20-APRIL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0093.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:

```
; LENGTH: 659 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: ORF 2, MEXICO, FIGURE 9
; US-07-870-985A-20

Alignment Scores:
Pred. No.: 0 Length: 659
Score: 659.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-07-870-985A-20 (1-659)

QY 1 ATGGCGCCTAGCGCTCTTTCTGCTGTGTTCTCTGTTCTGCTTATGTCGCCGCGCCA 60
Db 1 MetArgProArgProLeuLeuLeuLeuPheLeuLeuPheLeuProMetLeuProAlaPro 20
QY 61 CCGACCGGTGACGGCTCTGGCGCGCGCTGCTGGCGCGCGAGCGGGTACCGGGCGTGGT 120
Db 21 ProThrGlyGlnProSerGlyArgArgArgGlyArgArgSerGlyGlyThrGlyGlyGly 40
QY 121 TTCTGGGGTGACGGGGTGTATCTCAGCCCTTCGCAATCCCGCTATATTCATCCAAACCAAC 180
Db 41 PheTrpGlyAspArgValAspSerGlnProPheAlaIleProTyrIleHisProThrAsn 60
QY 181 CCCTTTGCCCGCAGAGTTGCCGCTGCGTCCGGTCTGAGCTCGCTTCGCTTCGCCCAACGACC 240
Db 61 ProPheAlaProAspValAlaAlaAlaSerGlySerGlyProArgLeuArgGlnProAla 80
QY 241 CGGCCACTTGGCTCCACATTGGCGAGATCAGGCCAGCGCCCTCCGCTGCTCCCGTGGC 300
Db 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
QY 301 CGACCTGCGCACAGCGGGGTGCGCGCTGACGGCTGAGCGCTGGCGCGCTGCCGATGACACCTCA 360
Db 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCCGTCCCGGAGGTGATTCTCGCGGTCAATCTACCGCCGACAGTATAATTGCTACT 420
Db 121 ProValProAspValAspSerArgGlyAlaIleLeuArgGlnTyrAsnLeuSerThr 140
QY 421 TCACCCCTGACATCCTCTGTGGCTCTGGCACTAATTTAGTCTGTATGCGAGCCCGCTT 480
Db 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuTyrAlaAlaProLeu 160
QY 481 AATCGCTCTGCGCTCGAGACGGTACTAATACTACATTATGGCCACAGAGCGCTCC 540
Db 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATATGACAGTACCGGGTTCGCCGCTACTATCCGTATCCGTACCGCCCTAGTGCTAAT 600
Db 181 AsnTyrAlaGlnTyrArgValAlaArgAlaThrIleArgTyrArgProLeuValProAsn 200
QY 601 GCAGTTGGAGGCTATGCTATATCCATTTCTTCTGGCTTCAACCAACCAACACCCCTACA 660
Db 201 AlaValGlyGlyTyrAlaIleSerIleSerPheThrProGlnThrThrThrProThr 220
QY 661 TCTGTGTACATGAATTCACATTAATCTCCACTGATGTGAGGATCTTTGTTCAACCTGGCATA 720
Db 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATTGGTCATCCCAAGCAGCGCTTCACTACCGCAATCAAGGTTGGCGCTCG 780
Db 241 AlaSerGluLeuValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSer 260
QY 781 GTTGAGACATCTGGTGTGCTGAGGAGGAGGACCTCGGTGCTTGTGATGTTATGTCATA 840
Db 841
```

```
Db 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCTCCAGTTAACTCCTATACCAATACCCCTTATACGGTGCCTTGGCTTACTG 900
Db 281 HisGlySerProValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCTTAGAGCTTGGATTTCCGAATCTCACCACTCTTAACCAATACACAGTGTG 960
Db 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTACTCCAGACTGTCTGTCATCTCCCGCCAGGGCGCGAGCTCGGAGCTG 1020
Db 321 SerArgTyrSerSerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeu 340
QY 1021 ACCCAACTGCAGCCACGAGTTTCATGAAGATCTCCACACTTACCGGCTTAATGGGTA 1080
Db 341 ThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyVal 360
QY 1081 GGTGAAGTCCGCGCGGATAGCTCTAACATTAACCTTGTGTGACACGCTCCTCGGC 1140
Db 361 GlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGly 380
QY 1141 GGGTCCCGACAGAAATTAATTTGCTGGCTCGCGGCGCAACTGTTTTATCCCGCCGGTT 1200
Db 381 GlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProVal 400
QY 1201 GTCTCAGCCTATGGCGAGCCCAACCGTGAAGCTCTATACATCAGTGGAGAATGCTCAGCAG 1260
Db 401 ValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGln 420
QY 1261 GATAAGGGTGTGTGTTATCCCGCCAGATATCGATCTTGTTGTTGTTGCTGCTGCTCATT 1320
Db 421 AspLysGlyValAlaIleProHisAspLeuAspLeuGlyAspSerArgValValIleGln 440
QY 1321 GATTATGACAAACAGCATGACAGAGATCGGCCCAACCGCTCGCTGCGCCATCTCGGCT 1380
Db 441 AspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArgPro 460
QY 1381 TTTTCTGTTCTCCGAGCAATGATGTTGTTGGTGTCTCCTCAGTCCAGCCGAGTATGAC 1440
Db 461 PheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaGlnTyrAsp 480
QY 1441 CAGTCCCACTTACGGGTCTCAACTGGCCCGGTTTATATCTCGGACAGCGTCACTTGGTG 1500
Db 481 GlnSerThrTyrGlySerSerThrGlyProValTyrIleSerAspSerValThrLeuVal 500
QY 1501 AATGTTGCGATGCGCGCAGCGCGTACCGCATGCTGTGCTGCTCAAAAGTCAACCTC 1560
Db 501 AsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeu 520
QY 1561 GACGGGCGGCCCTCCGACTGTTGAGCAATATTCGAAGACATCTTTGTGCTCCCGCTT 1620
Db 521 AspGlyArgProLeuProThrValGlnGlnTyrSerLysThrPhePheValLeuProLeu 540
QY 1621 CGTGGCAAGCTCTCCTTTTGGAGCGCGGCAACAAAGCAGAGTTATCCTTATATATAT 1680
Db 541 ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyr 560
QY 1681 AATACTACTGTAGTGACAGATCTTGATTGAAATGTCGCGGCGCATCGGCTGCCCAT 1740
Db 561 AsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIle 580
QY 1741 TCAACCTATACACAGGCTTGGGCGCGTCCGTCGCCATTTCTGCGCGCGCGGTTTG 1800
Db 581 SerThrTyrThrThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeu 600
QY 1801 GCTCCAGCCTCGCCCTCGCTCTGCTGAGGATACCTTTTGATTATCCGGGGCGGCGCAC 1860
Db 601 AlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHis 620
QY 1861 ACATTGTGATCTCTGCCCTGAATGCGCGCTTTAGGCTCCAGGCTGTGCTTCCAG 1920
Db 621 ThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGln 640
```

QY 1921 TCAACTGTGCTGAGCTCCAGCGCCCTTAAAGTTAAGTGGGTAAAACTCGGGAGTTG 1977
Db |||||||
641 SerThrValAlaGluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 659

RESULT 7

US-08-477-292-16
; Sequence 16, Application US/08477292
; Patent No. 6291641
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas
; APPLICANT: McAtee, Patrick
; APPLICANT: Yarbough, Patrice
; APPLICANT: Zhang, Yifan
; TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cathleen M. Desjardins, M.D.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: US
; ZIP: 94063

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,292
; FILING DATE:
; CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/327,952
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, Allan A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: G32P5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)369-9500
; TELEFAX: (415)368-0709
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Hepatitis E virus (Mexico strain)
; INDIVIDUAL ISOLATE: r62kDa, FIGURE 4
US-08-477-292-16

Alignment Scores:

Pred. No.: 0 Length: 548
Score: 548.00 Matches: 548
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 83.16% Indels: 0
DB: 3 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-08-477-292-16 (1-548)

QY 334 GCTGTGGCGCTGCCCATGACACCTCACCGTCCCGGAGTTGATCTCGCGGTGCAATT 393
Db |||||||
1 AlaValAlaProAlaHisAspThrSerProValProAspValAspSerArgGlyAlaIle 20
QY 394 CTACGCCGCCAGTATAATTGTCTACTTCACCCCTGACATCTCTGTGGCCTCTGGCACT 453
Db |||||||
21 LeuArgArgGlnTyrAsnLeuSerThrSerProLeuThrSerSerValAlaIleSerGlyThr 40

QY 454 AATTTAGTCTGTATGCAGCCCCCTTAATCCCGCTCTGCCGTGACGAGCGGTACTAAT 513
Db |||||||
41 AsnLeuValLeuTyrAlaAlaProLeuAsnProProLeuProLeuGlnAspGlyThrAsn 60
QY 514 ACTCACATTATGGCCACAGAGCGCTCCAATTATGCACAGTACCGGTTGCCCGCTACT 573
Db |||||||
61 ThrHisIleMetAlaThrGluAlaSerAsnTyrAlaGlnTyrArgValAlaAlaArgAlaThr 80
QY 574 ATCCGTTACGGCCCCCTAGTGCTTAATGCAGTTGGAGGCTGATGTATATCCATTCTTTC 633
Db |||||||
81 IleArgTyrArgProLeuValProAsnAlaValAlaGlyGlyTyrAlaIleSerIleSerPhe 100
QY 634 TGGCCTCAAAACAACAACCCCTACATCTGTGTGACATGAATTCCTACTTCCACTGAT 693
Db |||||||
101 TrpProGlnThrThrThrThrProThrSerValAspMetAsnSerIleThrSerThrAsp 120
QY 694 GTCAGGATCTTGTTCAACTGGCATAGCATCTGAATTCGTATCCCAAGCGAGCGCCTT 753
Db |||||||
121 ValArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTACCGCAATCAAGGTTGGCGCTCGGTGACACATCTGTTGCTGAGGAGGAGGCC 813
Db |||||||
141 HisTyrArgAsnGlnGlyTrpArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACCTCCGCTCTTGTTCATGTTATGCATACATGGCTCTCCAGTTAACTCCCTATACCAATACC 873
Db |||||||
161 ThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThr 180
QY 874 CTTTATACCGGTGCCCTTGCTTACTGGACTTTTGCCTTAGAGCTTAGAGTTGCAATCTC 933
Db |||||||
181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuGluPheArgAsnLeu 200
QY 934 ACCACCTGTAACACCAATACAGTGTGTCCTGCTACTCCAGCAGCTGCTGCTCACTCCGCC 993
Db |||||||
201 ThrThrCysAsnThrAsnThrArgValSerArgTyrSerSerThrAlaAlaHisSerAla 220
QY 994 CGAGGGCGCGACGGGACTCGGAGGTGACACCAACTGCAGCCACCAGCGTTTCATGAAGAT 1053
Db |||||||
221 ArgGlyAlaAspGlyThrAlaGluLeuThrThrThrAlaAlaThrArgPheMetLysAsp 240
QY 1054 CTCCTACTTACCGGCTTAATGGGTAGGTGAAGTCGCGCGGGAGTACTTAACATTAA 1113
Db |||||||
241 LeuHisPheThrGlyLeuAsnGlyValGlyGluValGlyArgGlyIleAlaLeuThrLeu 260
QY 1114 CTTAACTTGTGTCGACAGCTCTCCGCGGCTCCGACAGATTAATTTTCGTCGGCTGC 1173
Db |||||||
261 LeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAlaGly 280
QY 1174 GGGCAACTGTTTTATTCCCGCCCGTTGCTCAGCCAATGGCGAGCCACCGTGAAGCTC 1233
Db |||||||
281 GlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeu 300
QY 1234 TATACATGATGAGAAATGCTCAGCAGGATAGGGTGTGCTATCCCCACCATATCGAT 1293
Db |||||||
301 TyrThrSerValGluAsnAlaGlnGlnAspLysGlyValAlaIleProHisAspIleAsp 320
QY 1294 CTTGTGATTTCGCGTGTGCTATTTCAGGATTATGACACCCAGCATGAGCAGATCGGCC 1353
Db |||||||
321 LeuGlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgPro 340
QY 1354 ACCCGTCGCTCGCCCATCTCGGCTTTTCTGTCTCCGAGCAAAATGATGACTTTGG 1413
Db |||||||
341 ThrProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrP 360
QY 1414 CTGTCCCTCACTGACCGCGAGTATGACCACTCCACTTACGGTCTCACTCGCCGGTT 1473
Db |||||||
361 LeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProVal 380
QY 1474 TATATCTCGACAGCGCTGACTTTGCTGAATGTTGCGACTGGCGCGCAGCGCCGTAGCCCA 1533
Db |||||||
381 TyrIleSerAspSerValThrLeuValAsnValAlaThrGlyValGlnAlaValAlaArg 400
QY 1534 TCGCTTGATGCTCCAAAGTCAACCTCGAGCGGCGCCCTCCCGACTCTTGAGCAATAT 1593

401	Db	SexLeuAspTrpSerLysValThrLeuAspGlyArgProLeuProThrValGluGlnTyr	420
1594	Qy	TCCAAGACATCTTTGTGCTCCCCCTTCGTGGCAAGCTCTCCTTTTCGGAGGCGCGCAC	1653
421	Db	SerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThr	440
1654	Qy	ACAAAAGACGGTTATCCTTTATAATTAAATACTACTGCTAGTCACACAGATTCGATGAA	1713
441	Db	ThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeuIleGlu	460
1714	Qy	AATGCTGCCGGCCATCGGGTCGCCATTTCAACTATATCCACAGCTTGGGCGCGTCCG	1773
461	Db	AsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrThrArgLeuGlyAlaGlyPro	480
1774	Qy	GTCCGCTTCTTCGGGGCGGGTTTTGGCTCCACGCTCCGCCCTGGCTGCTCTGGAGAT	1833
481	Db	ValAlaIleSerAlaAlaIleValLeuAlaProArgSerAlaLeuAlaLeuLeuGluAsp	500
1834	Qy	ACTTTTGATTATCCGGGGCGGGCGCACACATTTCGCTGCTCCCTGAATGCCGCGCT	1893
501	Db	ThrPheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysArgAla	520
1894	Qy	TTAGGCCCTCCAGGGTTGTGCTTTCCAGTCAACTGTCGTGAGCTCCAGCGCTTTAAAGTT	1953
521	Db	LeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLysVal	540
1954	Qy	AAGTGGGTAAAACTCGGGAGTTG	1977
541	Db	LvsValGlyLvsThrArgGluLeu	548

RESULT 8

```

US-08-542-634-14
; Sequence 14, Application US/08542634
; Patent No. 6214970
; GENERAL INFORMATION:
; APPLICANT: Fuerst, Thomas R.
; APPLICANT: McAttee, C. Patrick
; APPLICANT: Yarbough, Patrice O.
; APPLICANT: Zhang, Yifan
; NUMBER OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR
; TITLE OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542,634
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0293.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEtical: NO
; ORIGINAL SOURCE:

```


QY 241 CGGCCACTTGCTCACTGTCGAGATCAGGCCAGCGCCCTCGCTGCTCCCTCCGTCG 300
Db |||||
QY 81 ArgProLeuGlySerThrTrpArgAspGlnAlaGlnArgProSerAlaAlaSerArgArg 100
Db |||||
QY 301 CGACCTGCCACAGCGGGGCTGCGCGCTGACGGCTGTGGCGCTGCCCATGACACCTCA 360
Db |||||
QY 101 ArgProAlaThrAlaGlyAlaAlaAlaLeuThrAlaValAlaProAlaHisAspThrSer 120
QY 361 CCCGTCGGAGCTGATTCTCGGGGTGCAATCTACGCCGCCAGTATAATTGCTACT 420
Db |||||
QY 121 ProValProAspValAspSerArgGlyAlaAlaLeuArgGlnTrpAsnLeuSerThr 140
QY 421 TCACCCCTGACATCTCTGTGGCTCTGGCACTAATTAGTCTGTATGACGCCGCCCTT 480
Db |||||
QY 141 SerProLeuThrSerSerValAlaSerGlyThrAsnLeuValLeuThrAlaAlaProLeu 160
QY 481 AATCCGCTCTGCGCTGACGAGCGGTACTAATCTACATATTATGGCCACAGAGCCCTCC 540
Db |||||
QY 161 AsnProProLeuProLeuGlnAspGlyThrAsnThrHisIleMetAlaThrGluAlaSer 180
QY 541 AATTATGCACAGTACCGGGTTCGCCGCTACTATCTGTTACCGGCCCTAGTGCCTAAT 600
Db |||||
QY 181 AsnTrpAlaGlnTrpArgValAlaAlaArgAlaThrIleArgTrpArgProLeuValProAsn 200
QY 601 GCAGTTCGAGGCTATGCTATATCAATCTCTTCTGGCTCAAAACACACACCCCTACA 660
Db |||||
QY 201 AlaValGlyGlyTrpAlaIleSerIleSerPheTrpProGlnThrThrThrProThr 220
QY 661 TCTGTTGACATGAATTCATTACTTCCACTGATGTCAGGATCTTGTTCACCTGGCATA 720
Db |||||
QY 221 SerValAspMetAsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIle 240
QY 721 GCATCTGAATGGTATCCCAAGGAGCGCTCTCACTACCGCAATCAAGTTGGCGCTCG 780
Db |||||
QY 241 AlaSerGluLeuValIleProSerGluArgLeuHisTrpArgAsnGlnGlyTrpArgSer 260
QY 781 GTTCAGACATCTGCTGTGTGTGAGGAGAACCCACTCCGCTCTGTCATGCTTATGATA 840
Db |||||
QY 261 ValGluThrSerGlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIle 280
QY 841 CATGGCTCTCCAGTTAACTCTCATACCAATACCCCTTATACCGTGGCTTGGCTTACTG 900
Db |||||
QY 281 HisGlySerProValAsnSerTrpThrAsnThrProTrpThrGlyAlaLeuGlyLeuLeu 300
QY 901 GACTTTGCTTAGAGCTTGAGTTGGCAATCTCACCTCTGCTGTAACCAATACAGGTGTG 960
Db |||||
QY 301 AspPheAlaLeuGluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgVal 320
QY 961 TCCCGTTTACTCCAGCACTGCTCGTCA---CTCCGCGCGAGGGCGGACGAGCTCGCGA 1016
Db |||||
QY 321 SerArgTrpSerSerThrAlaArgHisArgLeuArg-ArgGlyAlaAspGlyThrAlaG 340
QY 1017 GCTGACCACAACTGACGCCACCGGTTTCATGAAAGATCTCCACTTTACCGCCCTTAATGG 1076
Db |||||
QY 340 uLeuThrThrThrAlaAlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnG 360
QY 1077 GGTAGTGTAAGTCCGCGGAGTAGCTTAACATTACTTAACCTTGCTGACAGCTCCT 1136
Db |||||
QY 360 yValGlyGluValGlyArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLe 380
QY 1137 CGCGGGCTCCGACAGAAATTAATTTCTCGCTGGCGGCAACTGTTTATTCGCGGCC 1196
Db |||||
QY 380 uGlyGlyLeuProThrGluLeuIleSerSerAlaGlyGlyGlnLeuPheTrpSerArgPr 400
QY 1197 GGTGTCTTCAGCAATGGCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGATGCTCA 1256
Db |||||
QY 400 oValValSerAlaAsnGlyGluProThrValLysLeuTrpThrSerValGluAsnAlaG 420
QY 1257 GCAGATGAAGGTGTGTATCCGCCACAGATATCGATCTTGTGATTCGGTGTGGTCA 1316
Db |||||
QY 420 nGlnAspLysGlyValAlaIleProHisAspIleAspLeuGlyAspSerArgValVal 440
QY 1317 TCAGGATTATGACACCAACGATGACAGGATCGGCCACCGCTCGCTGCGCCATCTCG 1376

Db 440 eGlnAspTyAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerAr 460
QY 1377 GCCTTTTCTGTTCTCCGAGCAATGATGACTTTTGGCTGTCCTCTACTGACGCGGAGTA 1436
Db 460 gProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTy 480
QY 1437 TGACCACTGCACATTACGGTTCGTCACCTGGCCGGTTTATATCTCGGACACGCTGACTTT 1496
Db 480 rAspGlnSerThrTrpGlySerSerThrGlyProValTrpIleSerAspSerValThrLe 500
QY 1497 GGTCAATGTTGCGACTGCGCGCAGGCGGTAGCCCGATCGCTTGACTGGTCCAAAGTCAC 1556
Db 500 uValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValTh 520
QY 1557 CCTCGACGGGGCGCCCTCCGACTGTGTAGCAATATTTCCAGACATCTTTTGTGCTCCC 1616
Db 520 rLeuAspGlyArgProLeuProThrValGluGlnTrpSerLysThrPhePheValLeuPr 540
QY 1617 CTTTCGTGGCAAGCTCTCTTTTGGGAGCGGCACACAAAGCAGGTTATCTTATAA 1676
Db 540 oLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTrpTrpTrpAs 560
QY 1677 TTATAATACTACTGCTAGTCACCAAGATTCTGATTGAAAATGCTCCGCGCCATCGGTCGC 1736
Db 560 nTyAsnThrThrAlaSerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAl 580
QY 1737 CATTTCAACCTATACACCGAGCTTGGGCGCGCTCGCTCGCCATTTCTGCGCGCGCGT 1796
Db 580 aIleSerThrTrpThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaAlaVa 600
QY 1797 TTTGCTCCAGCTCCGCTCGCTCTGCTGAGAGTACTTTTGTATATCCGGGCGGCG 1856
Db 600 lLeuAlaProArgSerAlaLeuAlaLeuLeuGluAspThrPheAspTrpProGlyArgAl 620
QY 1857 GCACACATTTGATCACTTCTGCCCTGAATGCCGCGCTTAGGCTCCAGGGTTGTGCTTT 1916
Db 620 aHisThrPheAspAspPheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPh 640
QY 1917 CCAGTCAACTGCTGCTGAGTCCAGCGCTTAAAGTTAAGTGGGTAAACCTCGGAGATT 1976
Db 640 eGlnSerThrValAlaGluLeuGlnArgLeuLysValGlyLysThrArgGluLe 660
QY 1977 G 1977
Db 660 u 660
RESULT 10
US-09-462-606-48
; Sequence 48, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267U1
; CURRENT APPLICATION NUMBER: US/09/462,606
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-48
Alignment Scores: 0 Length: 660
Pred. No.:

APPLICANT: Bradley, Daniel W.
APPLICANT: Twu, Jr-Shin
APPLICANT: Purdy, Michael A.
APPLICANT: Tam, Albert W.
APPLICANT: Krawczynski, Krzysztof Z.
APPLICANT: Yarbough, Patrice D.
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/259,148A

FILING DATE: 13-JUN-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 822,335

FILING DATE: 17-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 505,888

FILING DATE: 05-APR-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 420,921

FILING DATE: 13-OCT-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 367,486

FILING DATE: 16-JUN-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 336,672

FILING DATE: 11-APR-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 208,997

FILING DATE: 17-JUN-1988

ATTORNEY/AGENT INFORMATION:

NAME: Sholtz, Charles K.

REGISTRATION NUMBER: 38,615

REFERENCE/DOCKET NUMBER: 4600-0093.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 435 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: C2, MEXICO, FIGURE 9

US-08-259-148A-18

Alignment Scores:

Pred. No.: 0 Length: 435

Score: 435.00 Matches: 435

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 66.01% Indels: 0

DBs: 1 Gaps: 0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-08-259-148A-18 (1-435)

QY 673 AATTCATTACTTCACGTGATCGAGGATCTTGTTCAACCTGGCATAGCAFTCGAATTG 732

DB 1 AsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGluLeu 20

QY 733 GTCATCCCAAGCGAGCGCCTTCACTACGCAATCAAGTTCGGCTCGCTTGGAGACATCT 792
DB 21 ValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThrSer 40
QY 793 GGTGTTGCTGAGGAGGAAGCCACCTCCGGTCTTCTCATGTTATGCATACATGGCTCTCCA 852
DB 41 GlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerPro 60
QY 853 GTTAACCTCTATACCAATACCCCTTATACCGGTGCCCTTGGCTTACTGGACTTTCGCTTA 912
DB 61 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 80
QY 913 GAGCTTGAGTTTCGCAATCTCACACCTGTAAACCAATACACGATGTCCTCCGTTACTCC 972
DB 81 GluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgValSerArgTyrSer 100
QY 973 AGCACTGCTGCTCACTCCGCCGAGGGCCGACGAGCTCGGAGCTGACCACTCA 1032
DB 101 SerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeuThrThrAla 120
QY 1033 GCCACCAAGTTCATGAAAGATCTCGACTTTACCGGCCCTTAATGGGTAGGTGAAGTCGC 1092
DB 121 AlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyValGlyGluValGly 140
QY 1093 CGCGGATAGCTTAAACATTACTTAACTTCTGACACGCTCTCGGGGGGCTCCCGACA 1152
DB 141 ArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThr 160
QY 1153 GAATTAATTCGTCGGCTGGCGGCACACTGTTTTATTCGCCCGGTTGTCTCAGCCCAAT 1212
DB 161 GluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsn 180
QY 1213 GCGAGCCCAACGCTGAGCTCTATACATCAGTGGAGATGCTCAGCAGATAGGGTGT 1272
DB 181 GlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyVal 200
QY 1273 GCTATCCCCCACGATATCGATCTTGGTGATTCGGGTGGTGCATTACAGATTATCACAAC 1332
DB 201 AlaIleProHisAspIleAspLeuGlyAspSerArgValIleGlnAspTyrAspAsn 220
QY 1333 CAGCATGACGAGGATCGGCCCAACCCCGCTCGCCATCTCGGCTTTTCTGTTCTC 1392
DB 221 GlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeu 240
QY 1393 CGAGCAATGATGATCTTGGCTGCTCCTCAGCAGCGCGATGATGACAGTCCACTTAC 1452
DB 241 ArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyr 260
QY 1453 GGTGCTCAACTGGCCCGGTTTATATCTCGACACGCTGACTTTGGTGAATGTTCCGACT 1512
DB 261 GlySerSerThrGlyProValTyrIleSerAspSerValThrLeuValAsnValAlaThr 280
QY 1513 GCGCGCAGCGCGTAGCCCGATCGCTTCAGCTGGTCCAAAGTCACCTCGAGGGCGGCC 1572
DB 281 GlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeuAspGlyArgPro 300
QY 1573 CTCGGCTGTTGAGCAATATCCAGACATCTTGTGCTCCCTTCCCTTCCGTCGAGAGCTC 1632
DB 301 LeuProThrValGluGlnTyrSerLysThrPhePheValLeuProLeuArgGlyLysLeu 320
QY 1633 TCCTTTTGGAGCGCGCACAAACAGCGTTATCTCTATAATTATAATACACTACTGCT 1692
DB 321 SerPheTrpGluAlaGlyThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAla 340
QY 1693 AGTGACCAAGATTCTGATTGAAAAATGCTCGCGGCCATCGGGTCGCCATTTCACCTATACC 1752
DB 341 SerAspGlnIleLeuIleGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThr 360
QY 1753 ACCAGCTTGGGCGCGGTCCGGTCGCCATTTCTGGCGCGCGGTTTGGCTCCAGCTCC 1812
DB 361 ThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeuAlaProArgSer 380

```
QY 1813 GCGCTGGCTCTGCTGAGGATACCTTTTGATTATCCGGGGCGGCGCACACATTGATGAC 1872
Db 381 AlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHisThrPheAspAsp 400
QY 1873 TTCTGCCCTGAATGCCGCTTTAGCCCTCCAGGGTTGCTTTCCAGTCAACTGTCGCT 1932
Db 401 PheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAla 420
QY 1933 GAGCTCCAGCGCTTAAGCTTAAGTGGCTAAACTCCGGAGTTG 1977
Db 421 GluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 435
RESULT 12
US-08-484-054-18
; Sequence 18, Application US/08484054
; Patent No. 5770689
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,054
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 870,985
; FILING DATE: 20-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 822,335
; FILING DATE: 17-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,888
; FILING DATE: 05-APRIL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 420,921
; FILING DATE: 13-OCTOBER-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 367,486
; FILING DATE: 16-JUNE-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: C2, MEXICO, FIGURE 9
US-08-484-054-18
Alignment Scores:
Pred. No.: 0 Length: 435
Score: 435.00 Matches: 435
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.01% Indels: 0
DB: 1 Gaps: 0
US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-08-484-054-18 (1-435)
QY 673 AATTCCATTACTTCCACTGATGTACAGATTCTTGTTCACACCTGGCATAGCATCTGAATTG 732
Db 1 AsnSerIleThrSerThrAspValAlaGlnLeuValGlnProGlyIleAlaSerGluLeu 20
QY 733 GTCATCCCAAGCGAGCGCTTCACTACCCCAATCAAGTTGCGCTCGGTGAGACATCT 792
Db 21 ValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTyrArgSerValGluThrSer 40
QY 793 GGTGTTGCTGAGGAGGAGCCACCTCCGCTCTTGTTCATGTTATGCATATACATGCTCTCCA 852
Db 41 GlyValAlaGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerPro 60
QY 853 GTTAACTCCTATACCAATACCCCTTATACCGGTGCGCTTGTGCTTACTGGACTTTGCTTAA 912
Db 61 ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu 80
QY 913 GAGCTTGAGTTTCGCAATCTCACACCTGTAAACCAATACAGTGTGTCCGTTACTCC 972
Db 81 GluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgValSerArgTyrSer 100
QY 973 AGCACTGCTGCTCACTCCGCGGAGGCGGAGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGCAG 1032
Db 101 SerThrAlaArgHisSerAlaArgGlyAlaAspGlyThrAlaGluLeuThrThrAla 120
QY 1033 GCCACCGAGTTTCATGAAAGATCTCCACTTTACCGGCTTAAATGGGTAGTGAAGTCGCGC 1092
Db 121 AlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyValGlyGluValGly 140
QY 1093 CCGGGGATAGCTTAACATTACTTAACCTTGTGTCGACACGCTCTCTCGGGGGCTCCGACA 1152
Db 141 ArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThr 160
QY 1153 GAAATTAATTTGCTCGGCTGGCGGCAACTGTTTTTATTCGCCGCCGCTTCTCAGCAAT 1212
Db 161 GluLeuIleSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsn 180
QY 1213 GCGAGCCAAACCGTGAAGCTCTATACATCAGTGGAGAATGCTCAGCAGGATAAGGTGTT 1272
Db 181 GlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnAlaAspLysGlyVal 200
QY 1273 GCTATCCCGCCAGCATCATGCTTGGTGTATTCGCTGGTGGTTCATTTCAGGATTATGACAAC 1332
Db 201 AlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsn 220
QY 1333 CAGCATGAGCAGGATCGGCGCCACCCGCTCGCTCGCCCATCTCGGCTTTTCTGTCTTC 1392
Db 221 GlnHisGluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeu 240
QY 1393 CGAGCAATGATGTTACTTTGGCTGTCCCTCACTGCGCGGAGTATGACAGTCCACTTAC 1452
Db 241 ArgAlaAsnAspValLeuTriLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyr 260
QY 1453 GGTCTGCTCAACTGGCGCGGTTTATATCTCGGACACGCTGACTTTGTTGATGTTGCGACT 1512
Db 261 GlySerSerThrGlyProValTyrIleSerAspSerValThrLeuValAsnValAlaThr 280
```

1513 GCGCGCAGCGCGTAGCCGATGCTTGAATGCTCAAAAGTCACCCCTCGACGGCGCGCC 1572
Db GlyAlaGlnAlaValAlaArgSerLeuAspTrpSerLysValThrLeuaspGlyArgPro 300
1573 CTCGCCAGCTGTGAGCAATATCCAAAGACATCTTTGTGCTCCCTTCGTTGGCAAGCTC 1632
Db LeuProThrValGluGlnTrpSerLysThrPheValLeuProLeuArgGlyLysLeu 320
1633 TCCTTTTGGAGCGCGGCAACAAAGCAGGTTATCCTTATAATTAATACTACTGCT 1692
Db SerPheTrpGluAlaGlyThrThrLysAlaGlyTrpProTyrAsnLysThrThrAla 340
1693 ATGACCAAGATCTTGATGAAATGTCGCGCCATCGGTCGCCATTTCAACCTATACC 1752
Db SerAspGlnLeuLeuIleGluAsnAlaAlaGlyHisArgValAlaIleSerThrThr 360
1753 ACCAGGCTGGCGCGGTCCGTCGCCATTTCTGCGCGCGCGGTTTGGCTCAGCTCC 1812
Db ThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeuAlaProArgSer 380
1813 GCGCTGGCTCTCTGAGGATACTTTTGATTATTCGGGGCGGCGCACACATTTTGATGAC 1872
Db AlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHisThrPheAsp 400
1873 TTCGCGCTGAATCGCGCTTTAGCCCTCCAGGTTGTGCTTCCAGTCACACTGTCGCT 1932
Db PheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAla 420
1933 GAGCTCCAGCGCTTAAAGTTAAGTGGTAAAGTGGTAAAGTGGGAGTTG 1977
Db GluLeuGlnArgLeuLysValLysValGlyLysThrArgGluLeu 435

RESULT 13

US-07-876-941A-18
; Sequence 18, Application US/07876941A

; Patent No. 5885768

; GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory R.

; APPLICANT: Bradley, Daniel W.

; APPLICANT: Tam, Albert W.

; APPLICANT: Mitchell, Carl

; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and

; TITLE OF INVENTION: Antibodies

; NUMBER OF SEQUENCES: 76

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates

; STREET: 350 Cambridge Avenue, Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/876,941A

; FILING DATE: 01-MAY-1992

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 822,335

; FILING DATE: 17-JAN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 505,888

; FILING DATE: 05-APRIL-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 420,921

; FILING DATE: 13-OCTOBER-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 367,486

; FILING DATE: 16-JUNE-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 336,672
; FILING DATE: 11-APRIL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 208,997
; FILING DATE: 17-JUNE-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0093.33
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: C2, MEXICO, FIGURE 9
; US-07-876-941A-18

Alignment Scores:

Pred. No.:	0	Length:	435
Score:	435.00	Matches:	435
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	66.01%	Indels:	0
DB:	2	Gaps:	0

US-09-851-410A-10_COPY_5117_7096 (1-1980) x US-07-876-941A-18 (1-435)

QY	673	AATTCATTACTTCCACTGATGTCAGGATTCCTGTTCAACCTGGCATAGCATCTGAATG	732
Db	1	AsnSerIleThrSerThrAspValArgIleLeuValGlnProGlyIleAlaSerGluLeu	20
QY	733	GTCATCCCAAGCGAGCGCTTCACCTACCGCAATCAAGTTGGCGCTCGGTGAGACATCT	792
Db	21	ValIleProSerGluArgLeuHisTyrArgAsnGlnGlyTrpArgSerValGluThrSer	40
QY	793	GCTGTGCTGAGGAGGAGCCACCTCCGCTCTTGTCATGTTATGTCATACATCGCTCTCCA	852
Db	41	GlyValAlaGluGluGluAlaThrSerGlyLeuValMetLeuCysIleHisGlySerPro	60
QY	853	GTTAACTCCTATACCAATACCCCTTATACCGGTGCCCTTGCTTACTGGACTTTCCCTTA	912
Db	61	ValAsnSerTyrThrAsnThrProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeu	80
QY	913	GAGCTTGAGTTTCGCAATCTCACCACCTGTAACCAATACACGTGTGTCGGTTACTCC	972
Db	81	GluLeuGluPheArgAsnLeuThrThrCysAsnThrAsnThrArgValSerArgTyrSer	100
QY	973	AGCACTGCTGCTCACTCCGCGCGGAGCGGCGGAGCTCGGAGCTGACCACTCACTCA	1032
Db	101	SerThrAlaArgHisSerAlaArgGlyAlaaspGlyThrAlaGluLeuThrThrAla	120
QY	1033	GCCACCGGTTTCATGAAAGATCTCCACTTTACCGGCTTAATGGGGTAGGTGAAGTCGGC	1092
Db	121	AlaThrArgPheMetLysAspLeuHisPheThrGlyLeuAsnGlyValGlyGluValGly	140
QY	1093	CGCGGATAGCTTAACATTAACCTTCCTGACACGCTCTCCGCGGGCTCCCGACA	1152
Db	141	ArgGlyIleAlaLeuThrLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThr	160
QY	1153	GAATTAATTTTCGTCGGCTGGCGGCAACTGTTTTATTCCCGCCCGTTGTCACGCAAT	1212
Db	161	GluLeuLeuSerSerAlaGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsn	180
QY	1213	GCGAGCCAACTGTAAGCTTATATCATCATGTGAGAAATGCTCAGCAGGATAAGGGTGT	1272
Db	181	GlyGluProThrValLysLeuTyrThrSerValGluAsnAlaGlnGlnAspLysGlyVal	200

Qy	1273	GCTATCCCCACGATATCGATCTTGGTGATTCGCGTGTGTCTATTCAGGATTTATGACAAC	13332
Db	201	AlaIleProHisAspIleAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsn	220
Qy	1333	CAGCATGAGCAGGATCGGCCACCCCGTCGCCCTGCGCANCTCGGCCCTTTTCTGTTCCTC	13992
Db	221	GlnHisGluGlnAspArgProThrProSerProIleProIleSerArgProPheSerValLeu	240
Qy	1393	CGACCAATGATGTACTTTTGGCTTCCTCTCACTCAGCCGAGTATGACCAAGTCCACTTAC	1452
Db	241	ArgAlaAsnAspValLeuTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyr	260
Qy	1453	GGGTGCTCAACTGCGCCGGTTTATCTCGACAGCGTGACTTTGGTGATGTTCGCACT	1512
Db	261	GlySerSerThrGlyProValTyrIleSerAspSerValThrLeuValAsnValAlaThr	280
Qy	1513	GGCGCGCAGGGCGGTAGCCCGATCCGCTGACTGTGTCCAAAGTCACCCTCGACGGCGGCC	1572
Db	281	GlyAlaGlnAlaValAlaArgSerLeuAspTrpSerIysValThrLeuAspGlyArgPro	300
Qy	1573	CTCCCGACTGTTGAGCAATATTCAGACATCTTTGTGTCTCCCGCTTCGTGGCAAGCTC	1632
Db	301	LeuProThrValGluGlnTyrSerIysThrPhePheValLeuProLeuArgGlyLysLeu	320
Qy	1633	TCCCTTTTGGGAGCGCGGCACAAACAGCAGTTATCCTTATAATTATATACTACTGCT	1692
Db	321	SerPheTrpGluAlaGlyThrThrIysAlaGlyTyrProTyrAsnTyrAsnThrThrAla	340
Qy	1693	AGTGACCAAGATCTGATWTGAAAAATGCTGCGCGCAATCGGTGCGCATTTCACTATACC	1752
Db	341	SerAspGlnIleLeuIleGluAsnAlaGlyHisArgValAlaIleSerThrTyrThr	360
Qy	1753	ACRAGGCTTGGGGCGGTCCGCTCGCCATTTCTCGCGCGGGTTCCTTCACCGCTCC	1812
Db	361	ThrArgLeuGlyAlaGlyProValAlaIleSerAlaAlaValLeuAlaProArgSer	380
Qy	1813	GCCCTGGCTCTGCTGGAGGATCTTTTGATTATCCGGGCGGGCGCACACATTTGATGAC	1872
Db	381	AlaLeuAlaLeuLeuGluAspThrPheAspTyrProGlyArgAlaHisThrPheAspAsp	400
Qy	1873	TTCTGCCCTGAAATCCGCGCTTTTAGGCCTCCAGGGTTGTCTTTCAGTCAACTGTCGT	1932
Db	401	PheCysProGluCysArgAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAla	420
Qy	1933	GAGCTCCACGCGCTTTAAAGTTTAAGGTGGGTAAAACTCGGGAGTTG	1977
Db	421	GlnLeuGlnArgLeuIleValIleValGlnGlyvsThrArgGluLeu	435

RESULT 14

```

US-07-870-985A-18
; Sequence 18, Application US/07870985A
; Patent No. 6455492
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory R.
; APPLICANT: Bradley, Daniel W.
; APPLICANT: Twu, Jr-Shin
; APPLICANT: Purdy, Michael A.
; APPLICANT: Tam, Albert W.
; APPLICANT: Krawczynski, Krzysztof Z.
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SCRAMBLER: Patent In. Release #1.0 Version #1.25

```



```
QY 694 GTGAGGATCTTGTTCACCTGGCATAGCATCTGAATTGGTCATCCCAAGCGAGCGGCTT 753
Db 121 VaArgIleLeuValGlnProGlyIleAlaSerGluLeuValIleProSerGluArgLeu 140
QY 754 CACTACCGCAATCAAGGTGGCGTTCGGTTCAGACATCTGGTGTGCTGAGGAGAACCC 813
Db 141 HisTyrArgAsnGlnGlyTyrArgSerValGluThrSerGlyValAlaGluGluAla 160
QY 814 ACTTCCGGTCTGTGCATGTATTGATCATATGGCTCTCCAGTTAACTCCTATACCATAACC 873
Db 161 ThrSerGlyLeuValMetLeuCysIleHisGlySerProValAsnSerTyrThrAsnThr 180
QY 874 CCTTATACCGGTGCGCTTGGCTTACTCGACTTTCCTTAGAGCTTGGTTCGCAATCTC 933
Db 181 ProTyrThrGlyAlaLeuGlyLeuLeuAspPheAlaLeuGluLeuPheArgAsnLeu 200
QY 934 ACCACCTGTAAACACCAATACACGCTGCTCCGCTTACTCCAGCAGCTGCTCGTCA---CTC 989
Db 201 ThrThrCysAsnThrAsnThrArgValSerArgTyrSerSerThrAlaArgHisArgLeu 220
QY 990 CGCCCGAGGGCGGCGGACTCGGAGCTGCACCACTGCACCACTGCACCACTGCAGTTCAATGAA 1049
Db 221 Arg-ArgGlyAlaAspGlyThrAlaGluLeuThrThrAlaAlaThrArgPheMetLeu 240
QY 1050 AGATCTCCACTTTTACCGGCTTAATGGGTAGGTAGTGGCGCGGAGTACTCTAAC 1109
Db 240 SasplLeuHisPheThrGlyLeuAsnGlyValGlyGluValGlyArgGlyIleAlaLeuTh 260
QY 1110 ATTACTTAACCTTGTGTGACACGCTCCTCGCGCGGCTCCCGACAGAAATTAATTCGTCGCG 1169
Db 260 rLeuLeuAsnLeuAlaAspThrLeuLeuGlyGlyLeuProThrGluLeuIleSerSerAl 280
QY 1170 TGGCGGGCACTGTTTATTCGCGCGGCTGCTCAGCCAAATGGCGAGCCAAACCGTGAA 1229
Db 280 aGlyGlyGlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrVally 300
QY 1230 GCTCTATACATCAGTGGAGATCTCAGCAGGATAGGGTGTGCTATCCCCCAGCATAT 1289
Db 300 sleuTyrThrSerValGluAsnAlaGlnAspLysGlyValAlaIleProHisAspI1 320
QY 1290 CGATCTTGTGTGATTCGCTGTGCTCATTGAGATTATGACCAACAGCAGCAGCAGGATCG 1349
Db 320 eAspLeuGlyAspSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspAr 340
QY 1350 GCCCACCCTGCTCGCTCGCCCATCTCGCGCTTTTCTGTTCTCCGAGCAAAATGATGACT 1409
Db 340 gProThrProSerProAlaProSerArgProPheSerValLeuAlaAsnAspValle 360
QY 1410 TTGGCTGTCCCTCACTGCGAGCGAGTATGACCACTTACGGTCCGCTCACTGGGCC 1469
Db 360 uTrpLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyPr 380
QY 1470 GGTTTATATCTGGACAGCGTGACTTTGTGTAATGTTGCGACTGGCGCGCAGCGGTAGC 1529
Db 380 oValTyrIleSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAl 400
QY 1530 CCGATCGCTTGACTGCTCAAAGTCAACCTCGACGGCGGCCCTCCCGACTGTTTGAGCA 1589
Db 400 aArgSerLeuAspTrpSerLysValThrLeuAspGlyArgProLeuProThrValGluGl 420
QY 1590 ATATTCCCAAGACATTTCTTGTGCTCCCTTGTGCTGCGAGCTCTCTTTTGGAGCGCG 1649
Db 420 nTyrSerLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGl 440
QY 1650 CACAACAAAGAGGTTATCCCTATATTAATACTACTGCTAGTACCAGATTCTGAT 1709
Db 440 yThrThrLysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnIleLeuI1 460
QY 1710 TGAAATGCTCGCGCATCGGCTCGCATTTCAACCTATACCAGCGCTTGGGCGCG 1769
Db 460 eGluAsnAlaAlaGlyHisArgValAlaIleSerThrTyrThrThrArgLeuGlyAlaGl 480
QY 1770 TCCGCTCGCCCATTTCTGCGGCGCGGCTTTGGCTCCAGCTCCGCGCTGGCTCTGCTGA 1829
```

```
Db 480 yProValAlaIleSerAlaAlaAlaValLeuAlaProArgSerAlaLeuAlaLeuGl 500
QY 1830 GGATACTTTTGAATTATCCGGGCGGGCGGCACACATTTTGATGACTTCTGCCCTGAATGCCG 1889
Db 500 uAspThrPheAspTyrProGlyArgAlaHisThrPheAspAspPheCysProGluCysAr 520
QY 1890 CGCTTTAGGCTCCAGGTTGTGCTTTCCAGTCAACTGTGCTGAGCTCCAGCGCGCTTAA 1949
Db 520 gAlaLeuGlyLeuGlnGlyCysAlaPheGlnSerThrValAlaGluLeuGlnArgLeuLy 540
QY 1950 AGTTAAGTGGGTAAAACTCGGGAGTTG 1977
Db 540 sValLysValGlyLysThrArgGluLeu 549
```

Search completed: August 16, 2004, 14:00:12
Job time : 52.9849 secs